Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

The Significance of the Contributions of A.N.Prior and Jerzy o in the Early History of Modern Temporal Logic

Øhrstrøm, Peter

Published in:
Logic and Philosophy of Time

Creative Commons License
CC BY-NC-ND 4.0

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

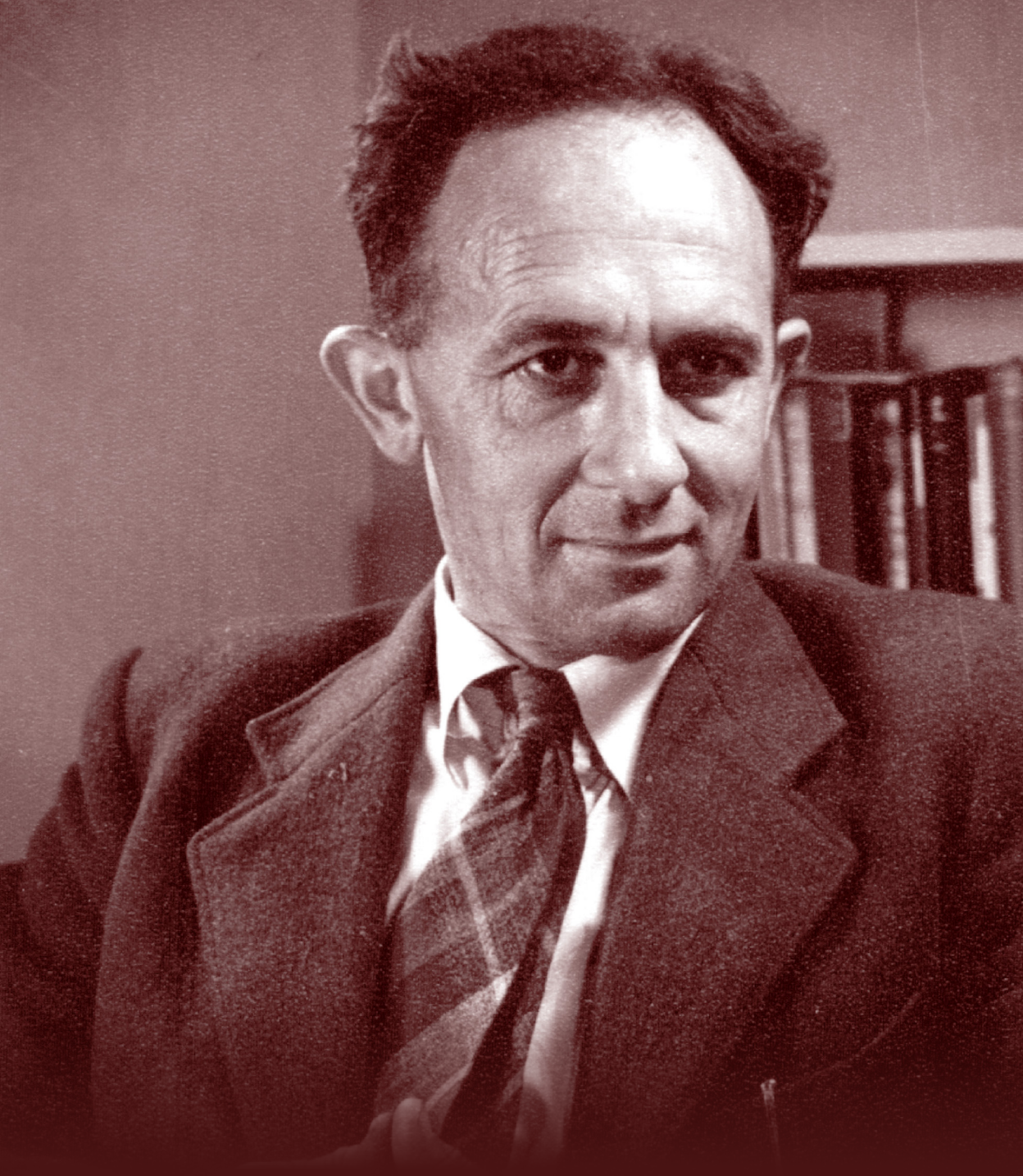
Link to publication from Aalborg University

Citation for published version (APA):
Øhrstrøm, P. (2019). The Significance of the Contributions of A.N.Prior and Jerzy o in the Early History of Modern Temporal Logic. In P. Blackburn, P. Hasle , & P. Øhrstrøm (Eds.), *Logic and Philosophy of Time: Further Themes from Prior, Volume 2* (1 ed.). Aalborg Universitetsforlag. Logic and Philosophy of Time Vol. 2

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.



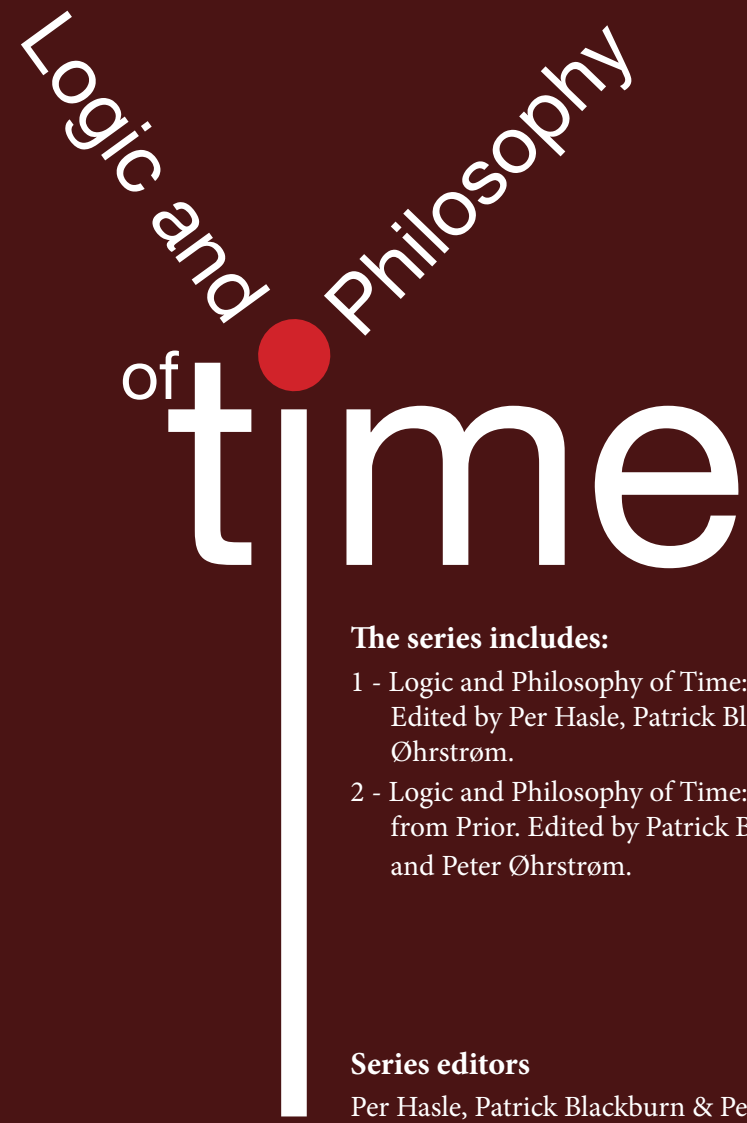
Logic and Philosophy of Time

Futher Themes from Prior, Vol. 2

Patrick Blackburn, Per Hasle and Peter Øhrstrøm (Eds.)

Logic and Philosophy of Time

A.N. Prior (1914-69) in the course of the 1950s and 1960s founded a new and revolutionary paradigm in philosophy and logic. Its most central feature is the preoccupation with time and the development of the logic of time. However, this was inseparably interwoven with fundamental questions about human freedom, ethics, and existence. This remarkable integration of themes also embodies an original and in fact revolutionary conception of logic. The book series, *Logic and Philosophy of Time*, is dedicated to a deep investigation and also the further development of Prior's paradigm.



The series includes:

1 - Logic and Philosophy of Time: Themes from Prior. Edited by Per Hasle, Patrick Blackburn, and Peter Øhrstrøm.

2 - Logic and Philosophy of Time: Further Themes from Prior. Edited by Patrick Blackburn, Per Hasle, and Peter Øhrstrøm.

Series editors

Per Hasle, Patrick Blackburn & Peter Øhrstrøm

Logic and Philosophy of Time:

Further Themes from Prior

Edited by:
Patrick Blackburn, Per Hasle, and
Peter Øhrstrøm

Logic and Philosophy of Time, Volume 2

AALBORG UNIVERSITY PRESS

Logic and Philosophy of Time: Further Themes from Prior
Edited by Patrick Blackburn, Per Hasle & Peter Øhrstrøm

1. Open Access Edition

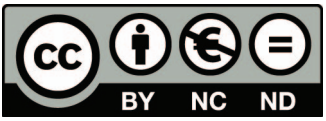
© The authors and Aalborg University Press, 2019

Interior design: Ulrik Sandborg-Petersen & Fatima Sabir
Cover design: akila by Kirsten Bach Larsen
Photo on front cover: A.N. Prior, courtesy of Martin Prior
Copy editing: Fatima Sabir
Set with the TeX Gyre Pagella & TeX Gyre Heros fonts

ISBN: 978-87-7210-265-8
ISSN: 2596-4372

This book is financially supported by the Danish Council for Independent Research | Culture and Communication

Published by:
Aalborg University Press
Langagervej 2
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk



Attribution-NonCommercial-NoDerivatives

Table of Contents

Preface: Nows and Thens in the Prior Project

Patrick Blackburn
Roskilde University, Denmark
patrick.rowan.blackburn@gmail.com

This is the second volume of a collection of papers centered on the work of Arthur Prior. It draws on material presented at two conferences: one held in Copenhagen from 22nd-24th November 2017, and a shorter event held at Roskilde University on 2nd March 2018 as part of the Annual Meeting of the Danish Philosophical Society.[1]

Like its predecessor, this collection is broad in sweep: it contains papers discussing the history of Arthur Prior's life and work, papers on philosophical themes introduced, elaborated, or alluded to by Prior, and papers that draw their inspiration from Arthur Prior's rich and varied contributions to logic. This subdivision is hardly sharp. Many contributions — much in the spirit of Prior's own work — move easily between history, philosophy, and logic. Moreover, like its predecessor, this volume concludes with a paper that views the Prior Internet Resources from the perspective of the digital humanities. In short, both volumes illustrate the breadth of the Prior Project, the topic to which we now turn.

"Prior Project" is a usefully ambiguous term. It can be read as referring to aspects of Arthur Prior's own lifelong intellectual project, an

[1] The first volume, *Logic and Philosophy of Time: Themes from Prior*, edited by Per Hasle, Patrick Blackburn and Peter Øhrstrøm, drew on the same November 2017 event in Copenhagen and from a still earlier conference held in Skagen (at the northernmost tip of Jutland) held from 30th May until 1st June 2017. The papers from the previous volume are listed, with page references, at the end of the preface.

exploration which wove his life and work tightly together. For example, Prior wrestled with theological issues (such as predestination) and these fed back into his writing (in the case of predestination, via his work on branching time). However "Prior Project" also has a more down-to-earth use: as shorthand for a project funded by the Danish Council for Independent Research called *The Primacy of Tense — A.N. Prior Now and Then*.

The project began in October 2016 and will run until at least October 2019. Headed by Peter Øhrstrøm, it brings together researchers, all actively involved with various aspects of Prior's work, from six Danish universities: Aalborg University, Aarhus University, Roskilde University, the University of Copenhagen, the University of Southern Denmark, and the Technical University of Denmark.

This prompts an obvious question: how has Denmark become a centre for research on Arthur Prior? Prior is perhaps New Zealand's best known philosophical logician,[2] but he is hardly a household name. Why the continued interest, and why so far from New Zealand?

Part of the answer lies in the richness of Prior's work. It is alive with ideas about what logic is, does, or might be, and how it links with philosophy, science, and theology. Much of this work was overlooked, perhaps forgotten, after Prior's death in 1969. A classic example is hybrid logic, which lay buried in plain sight in the pages of *Past, Present and Future* and *Papers on Time and Tense* until its first independent reinvention (nearly two decades after Prior died) by the Sofia School logicians Solomon Passy, Tinko Tinchev, George Gargov, and Valentin Goranko; only in the late 1980s was the link with Prior's earlier work made.

Another part of the answer, and the link with Denmark, lies in the efforts of Peter Øhrstrøm and Per Hasle. The publication of their joint 1995 book *Temporal Logic: From Ancient Ideas to Artificial Intelligence*, in which Prior's ideas featured prominently, helped rekindle interest in his work. Crucially, Øhrstrøm and Hasle also contacted Prior's widow, Mary Prior. This led to multiple trips to Oxford (Prior's last home),

[2]Richard Sylvan (*né* Routley) is the most obvious reason for inserting a qualifying 'perhaps' here. These two Kiwi logicians had much in common: both were born in the lower North Island of New Zealand, both were charismatic teachers and mentors, both married twice, both invigorated philosophical logic with new lines of inquiry which continue to be explored to this day, both spent their most productive years outside New Zealand, and both died abroad. Moreover, Sylvan's longstanding environmental activism curiously mirrors Prior's lifelong wrestling match with religion.

to years of work photographing and classifying Prior's *Nachlass* material, to an ongoing transcription effort (in which Martin Prior, Arthur Prior's son, continues to play an important role), to the organization of the 2014 Prior Centenary Meeting at Balliol College, and, most recently, to the Prior Project.

The project yokes together two strands of work: the "Now" aspect and the "Then" aspect. The "Now" aspect treats Prior's work as a source of philosophical and logical inspiration, and views Prior's life and work as fruitful objects of study. This aspect of the Prior Project is well reflected in this volume and its predecessor.

The "Then" aspect, however, does not fit so neatly between the pages of a book. Rather, the heart of the "Then" aspect is contained in 29 boxes in the Bodleian Library, Oxford, with the remainder spread over known and yet-to-be-discovered locations around the globe. For the "Then" aspect is the discovery, classification, curation, digitization, and distribution of Prior's *Nachlass*: draft books and papers, his off-the-cuff notes and textual annotations, the letters he sent and the letters he received. The cornerstones of the "Then" aspect are the online transcribers and the transcription workshops held periodically at Aalborg University — these make Prior's unpublished writings readable and available. For more on the ongoing attempts to transform Prior's physical *Nachlass* into a virtual *Nachlass*, see the final paper in this volume.

Time to turn to the volume's individual contributions. Their ordering mirrors that of the predecessor volume: loosely speaking, it begins with the more historically oriented papers, moves through more philosophically oriented papers to those with logical themes, and concludes with a view from digital humanities. Here is the listing:

The Return of Medieval Logic in the Philosophy of Time
David Jakobsen

Jakobsen argues that Prior's invention of tense logic marks a return of medieval logic to the philosophy of time. Drawing on material from the Prior archive at the Bodleian Library in Oxford, Prior's 1958 paper *The Syntax of Time Distinctions* is read as a solution to a problem discussed by Strawson and Quine in 1953. Jakobsen argues that Prior is here "inviting metaphysics into philosophy", thereby abandoning views that dom-

inated analytic philosophy in the first half of the twentieth century.

The Significance of the Contributions of A.N. Prior and Jerzy Łoś in the Early History of Temporal Logic

Peter Øhrstrøm and Per Hasle

Prior is usually regarded as the founder of modern temporal logic, but a recent paper by Marcin Tkaczyk and Tomasz Jarmużek argues that Jerzy Łoś deserves the title. Here Øhrstrøm and Hasle present the case for Prior. This is a topic we are unlikely to have heard the last of.

B-theory and Time Biases

Sayid R. Bnefsi

In his well-known paper *Thank Goodness That's Over*, Prior pointed out that we care not only about *what* experiences we have, but also about *when* we have them. But why, Prior asks, should a B-theorist care about timing? Here Bnefsi counters this "infamous kind of argument" against B-theory. Appealing to four-dimensionalism and an evaluative principle concerning whose experiences matter, he argues that B-theorists can justify some time-biased preferences.

Presentism and Cross-Time Relations

R. D. Ingthorsson

If only the present exists, relations with other times can't exist either. This "problem of cross-time relations" is a strong objection to presentism, the form of A-theory that Prior favoured. Ingthorsson here offers a robust defence of presentism. The Aristotelian view of causation and persistence does not, he argues, invoke relations between entities that exist at different times; indeed, it excludes the possibility of such relations. Moreover, rejecting the existence of the past does not lead to absurdity.

Ingthorsson, McTaggart's Paradox and the R-theory of Time
L. Nathan Oaklander

Ingthorsson recently argued that McTaggart's argument for the unreality of time rests on the principle of temporal parity: all times exist equally in a sense that is compatible with their being successive. Hence, as this principle underlies B-theoretic accounts of time, McTaggart's argument undermines B-theory. Oaklander counters with an appeal to Russell's account of time, arguing that R-theory is immune to objections based on temporal parity. He also argues that Ingthorsson's version of presentism is susceptible to a novel interpretation of McTaggart's argument.

Dummett on McTaggart's Proof of the Unreality of Time
Brian Garrett

Michael Dummett (in collaboration with E. J. Lemmon) was an early contributor to the technical development of tense logic; Prior knew their joint work and made use of it in *Past, Present and Future*. However, Dummett also proposed an ingenious analysis of McTaggart's proof of the unreality of time. Garrett here argues that Dummett's analysis is best viewed as a transformation of McTaggart's original argument into an argument showing the falsity of the observer-independence thesis.

A Logical Framework for the Spotlight Theory of Time
Ciro De Florio, Alessandro Giordani and Aldo Frigerio

Here the starting point is Broad's moving spotlight conception of time. De Florio, Giordani, and Frigerio first provide a conceptual analysis, highlighting its underlying metaphysical assumptions. They then introduce a logical language containing the standard Priorean tense operators, a stronger pair of tense operators, and a "now" operator. This they use to explore the philosophical background of Broad's theory, and to formalize such notions as "being a lost possibility".

Now There Will Be Trouble

Giuseppe Spolaore and Fabio Del Prete

Here we begin with some puzzling linguistic examples: sentences with "now" in initial position that differ in meaning from (otherwise identical) sentences with "now" in final position. As Spolaore and Del Prete show, these indicate an interaction between tense logic and pragmatics. This leads them to introduce a notion of "modal forcing", which provides a uniform account of both readings of "now", and lets them solve some modal-temporal problems concerning fatalism and the possibility of a changing past.

Actuality and Possibility in Branching Time: The Roots of Transition Semantics

Antje Rumberg

This paper examines the foundations of branching time semantics and addresses both technical and philosophical themes. Prior explored two approaches to branching time, but the Peircean approach he favoured is expressively weak, and the stronger Ockhamist approach introduces a global dimension into an otherwise local logic. Using a series of diagrams, Rumberg explicates an alternative: transition semantics. Her approach is strongly Priorean in spirit: it provides a *local* approach to branching time. Furthermore, it provides a setting in which the Peircean and Ockhamist options emerge as polar opposites in the spectrum of transition possibilities.

TRL Semantics and Burgess' Formula

Roberto Ciuni and Carlo Proietti

Like the previous one, this paper explores the foundations of branching time semantics. Under the microscope here is a variant of Prior's Ockhamist semantics, a Thin Red Line (TRL) semantics introduced by Bräuner, Øhrstrøm, and Hasle. Ciuni and Proietti show that this TRL semantics invalidates Burgess' formula $\Box G(\phi \rightarrow \Diamond F\varphi) \rightarrow \Diamond G(\varphi \rightarrow F\varphi)$, which *is* valid in Prior's Ockhamist semantics. The authors argue that this is unsatisfactory, and discuss how best to amend the proposed semantics to restore the validity of Burgess' formula (and two others).

The Role of Time in Phronetic Activities

Anne Gerdes

Time, Gerdes argues, is fundamental to the role of judgement in *phronetic* activities: it is crucial to the exercise of judgment and the evaluation of the moral worth of past actions. Furthermore, its importance is best explicated in terms of branching time semantics. Building on this, Gerdes suggests that an artificial ethical agent could never be strongly *phronimos*, as it would lack the ability to experience (and hence learn) from the relation between time and risk.

The Prior eAchive as Virtual Research Environment: towards Serendipity and Explorability

Fatima Sabir and Volkmar Poul Engerer

The "Then" component of the Prior Project is the ongoing attempt to transform Prior's physical *Nachlass* into a virtual *Nachlass*. But what has been done so far, and what should be done? The closing paper answers both questions. Sabir and Engerer first describe the history and current state of the Prior Internet Resources (PIR), and then argue that the PIR should be viewed as a Virtual Research Environment and developed in a way that will encourage explorability and serendipitous research.

TABLE OF CONTENTS OF THE PREVIOUS VOLUME

A Note on the Expressive Power of Temporal Discourse in Natural Language
David Rey, pages 103–116.

Counterfactuals, Causal Independence and Determinism
Lars Bo Gundersen, pages 117–122.

The Coat Problem. Counterfactuals, Truth-makers, and Temporal Specification
Francesco Gallina and Giuseppe Spolaore, pages 123–139.

Kant on Time and Change: A Series, B Series, or Both?
Hope Sample, pages 141–150.

Persistence Reconsidered: Beyond the Endurance/Perdurance Distinction
Florian Fischer, pages 151–166.

The Dynamic Present
Antony Galton, pages 167–187.

Tense and Omniscience
Ciro de Florio, Aldo Frigerio and Alessandro Giordani, pages 189–205.

The Truth about Osmo
Elton Marques, pages 207–223.

The Prior Internet Resources 2017: Information Systems and Development Perspectives
Volkmar Paul Engerer and Jørgen Albretsen, pages 225–251.

Acknowledgements

The editors are extremely grateful to Fatima Sabir for her seemingly tireless copy editing and bibliography checking, to Ulrik Sandborg-Petersen for his skillful transformation of a collection of Word and LaTeX files into a beautifully formatted whole, and to Kirsten Bach Larsen for the elegant cover designs of both volumes.

Bibliography

[1] Hasle, P., Blackburn, P. and Øhrstøm, P. (2018). *Logic and Philosophy of Time: Themes from Prior*. ISBN: 978-87-7210-009-8, ISSN: 2596-4364. Aalborg: Aalborg University Press.

[2] Øhrstøm, P. and Hasle, P. (1995). *Temporal Logic: From Ancient Ideas to Artificial Intelligence*. Berlin: Springer Verlag.

[3] Prior, A.N. (1967). *Past, Present and Future*. Oxford: Oxford University Press.

[4] Prior, A.N. (1968). *Papers on Time and Tense*. Oxford: Oxford University Press.

The Return of Medieval Logic in the Philosophy of Time

David Jakobsen
Aalborg University, Denmark
davker@hum.aau.dk

Abstract

This paper argues that A.N. Prior's invention of tense-logic constitutes a return of medieval logic in the philosophy of time. The argument proceeds from an analysis of W.V. Quine and P.W. Strawson's 1953 discussion about the inability of formal logic to analyse the tenses of ordinary language. Recent discoveries in the A.N. Prior archive at the Bodleian Library in Oxford reveal that J.J.C. Smart, in a letter to Prior, brought up their discussion in his rejection of Prior's invention of tense-logic. The correspondence will be discussed in terms of the importance of Prior's discovery of tense-logic and the presentation of this as a solution to the problem discussed by Strawson and Quine in *The Syntax of Time Distinctions* (1958, [9]). Recent discoveries reveal a close connection between this discussion and Prior's discussion of future contingency in *Diodoran Modality* (1955, [7]). These discoveries support the conclusion of this paper, which argues that Quine's insistence that modern logic must be conceived as tenseless was demonstrated by Prior to be false. Furthermore, through Smart's rejection in his correspondence with Prior, it will be shown that Smart persistently warned Prior not to present tense-logic at the John Locke Lectures in 1956. Apparently, Smart was convinced that Oxford would not appreciate Prior's logic and that Prior was wrong. Despite Smart's warnings, Prior's tense-logic was well received at Oxford. It is here argued that, to understand the importance of A.N. Prior's invention of tense-logic, as well as Quine's insistence that modern logic must be tenseless and Smart's warning against presenting it at Oxford, we must appreciate it as a return of medieval logic in the philosophy of time. This

return presented a challenge to the views of Quine and Smart—a challenge that can only be overcome by inviting metaphysics into philosophy, thereby abandoning the view of the analysis that dominated philosophy in the first half of the twentieth century.

Keywords: Philosophy of time, Medieval Logic, A.N. Prior, Strawson, Quine, Tenses, Tense-logic, Formal Logic.

1 Introduction

The advent of analytic philosophy at the end of the nineteenth and beginning of the twentieth century was, according to Bertrand Russell, the beginning of a "new philosophy", with the formal analysis of language at its centre (Russell 1959, p. 42, [16]). The formal analysis of language means, as Russell puts it, that "all sound philosophy should begin with an analysis of propositions', a truth that to him was 'too evident, perhaps, to demand a proof" (Russell 1992, p. 8, [17]). Russell's analysis has a purifying form that does not try to solve a philosophical problem, in a metaphysical sense, but rather to rid language of philosophical problems by demonstrating that they are not really problems after all:

> Every philosophical problem when it is subjected to the necessary analysis and purification, is found to be not really philosophical at all, or else to be, in the sense in which we are using the word, logical. (Russell 1914, p. 33, [15])

The new philosophy, or modern analytic philosophy, was therefore born with very little interest in medieval logic. For Russell, "logic in the Middle Ages, and down to the present day in teaching, meant no more than a scholastic collection of technical terms and rules of syllogistic inference. Aristotle had spoken, and it was the part of humbler men merely to repeat the lesson after him" (Russell 1914, p. 42, [15]).

Russell was only half right in his characterization of the way in which medieval philosophers such as Ockham and Scotus read Aristotle. Indeed, they had an immense degree of respect for The Philosopher, but they had no problem correcting him where they found him wrong. The works of medieval logicians differed from Russell's, not merely in their respect for Aristotle but also in the role philosophical problems played

in the development of logic. While the work of theologians, such as Ockham and Scotus played a role in setting the bounds of Orthodoxy, they were themselves potential subjects of heresy trials for unorthodoxy. The logical analysis of propositions was part of the larger goal of remaining within the bounds of orthodoxy, especially with regard to the relationship of God to time and modality (see Uckelman 2009, [21]). Of these issues, as Øhrstrøm and Hasle point out, "the most important question [...] was the problem of the contingent future" (Øhrstrøm and Hasle 1995, p. 87, [22]). The advances made in modal logic by the medieval logicians occurred in the context of attempting to solve problems concerning God, time and necessity (Hintikka and Spade 2017, [2]). The basic difference between Russell and medieval logicians such as Ockham and Scotus pertains to their view on analysis—what it is about and what it is for. Russell, like Ockham and Scotus, ascribes a purifying role to analysis, but where Ockham and Scotus aim at purifying Aristotle to get rid of false (i.e., unorthodox) philosophy, Russell aims at purifying language so that it does not lead to a conclusion that science would never be able to verify. For Ockham and Scotus, logic is an intrinsic part of true propositions, and analysis therefore aims at discovering whether different statements about the future or necessity align with Christian orthodoxy. Russell was, for this reason, perhaps wrong in another sense in calling his philosophy "a new philosophy". Russell's view on analysis was a natural development of the methodological role logic was assigned in modern times. Øhrstrøm and Hasle argue, with regard to the role of time reference, that a fundamental modern shift away from medieval logic began with Francis Bacon's emphasis on logic as a methodology and Descartes' view of mathematics as the model of all science. This, they argue, "would eventually lead to the dissociation of logic from language, that very connection which in Scholastic times had inter alia legitimized the study of propositions with time reference" (Øhrstrøm and Hasle 1995, p. 114, [22]). Thus, at the turn of the twentieth century, the scene looked bleak for medieval logic in terms of the philosophy of time. The incredible achievements made in formal logic seemed to be able to solve all the problems the medieval logicians struggled with by dissolving them into pseudo-problems. Things turned out differently, however. Contrary to Russell's expectations, the project of formalising language did not dissolve problems pertaining to the philosophy of time. The story of this development begins with philosophers

for whom the question of the nature of time is essential for an adequate analysis of tensed propositions. It is also a story of the return of medieval logic in the philosophy of time.

2 Ordinary Language and Tensed Propositions

The first half of the twentieth century was dominated by Russell's understanding of analysis. Despite its formal advantages, philosophers within the ordinary language tradition, such as P.F. Strawson, saw it is as a problem that formal analysis could not make sense of tensed propositions and asked, "how are we to express in the symbolism of quantifiers the difference between 'There is', 'There was', 'There will be?'" (Strawson 2011, p. 151, [20]). By way of example, Strawson considers the tensed statement

(1) There was at least one woman among the survivors.

He suggests that, in formal logic, (1) would take the form

(2) $\exists x : (x \text{ is a woman} \wedge x \text{ was among the survivors})$.

Given a direct translation of the form $\exists x$, the logical form of the statement is as follows:

(3) There is at least one person who is a woman and was among the survivors.

Contrary to (1), statement (3) suggests that the person is alive at the time of the utterance of the proposition. Since this assumption is not in (1), statement (3) cannot be a good analysis of it. Neither will it work to change the second "is" to "was", since this will only suggest that x somehow ceased to remain a woman:

(4) There is a least one person who was a woman and was among the survivors.

It is evident that such an analysis is hard pressed to find a way to make purely logical sense of the tensed verb "was". If (1) is a purely logical proposition, and thus does not refer to any singular fact—which is required for the statement to be merely about language—then "was" must be reduced to some symbol that turns the proposition into a tautology. Strawson's examples point out the absurdities of Russellian analysis. The final solution considered by Strawson is to maintain that the

existential quantifier must be interpreted as tenseless, such that "when we speak of persons and incidents the question of time reference does not arise" (Strawson 2011, p. 150, [20]). Aware of the unsatisfying nature of such a solution, Strawson says, "Nor can the difficulty be evaded by declaring '$\exists x$' in this sentence to be timeless; it is not true that when we speak of persons and incidents the question of time-reference does not arise" (Strawson 2011, p. 151, [20]).

Strawson thus concludes that formal logic is limited in its ability to analyse natural language sentences even though it is quite capable of making progress in analysing mathematical sentences. This strikes at the heart of the new philosophy's two principles by raising the question: how can the 'new philosophy' solve philosophical problems in Russell's sense, through the analysis of propositions, if it cannot adequately formalise propositions in which references to time are essential?

QUINE'S RESPONSE

Strawson had put his finger on a sore problem, and Quine takes up Strawson's example in *Mr. Strawson on Logical Theory* (1953, [14]). Here, he suggests the following way of rendering Strawson's example:

(5) $\exists x : (x$ [is] a woman $\wedge\, x$ was among the survivors$)$.

where both $\exists x$ (as Quine always adds) and [is] are understood in a timeless sense. It is important for Quine that the "only tenable attitude toward quantifiers and other notations of modern logic is to construe them always, in all context, as timeless" (Quine 1953, p. 442, [14]). "Was" should thus, according to Quine, be understood as involving a "reference presumably to some time or occasion implicit in the missing context" (Quine 1953, p. 442 [14]). Quine assumes that the context is given by a variable for some event, D, such that the proper proposition is

(6) $\exists x : (x$ [is] a woman $\wedge\, x$ [is] among the survivors of $D)$.

Unlike Strawson, Quine is not troubled by the evident quantification over "thing-events, four dimensional beings in space-time" (Quine 1953, p. 442, [14]). The various context-dependent facts of the values under quantification are just as timeless as is the universe of discourse when the values take on mathematical properties. Strawson's fundamental objection is that formal logic presents a severely limited picture of reality that is vastly different from the way reality is revealed to us in natural

language discourse. Quine's answer is that this is something we must live with if we truly appreciate the value of modern formal logic:

> The four dimensional view of space-time is part and parcel of the use of modern formal logic, and in particular the use of quantification theory, in application to temporal affairs. It may be felt to be a criticism of modern logic that it calls for so drastic a departure from the time-slanted IndoEuropean language structure. But the better way of looking at the matter is to recognize both in the four-dimensional approach, with its notable technical advantages, and in quantification theory, with its notable technical advantages, two interrelated contributions to scientific method.
>
> (Quine 1953, p. 442, [14])

The key phrase to accept in Quine's departure from the natural language of discourse is "technical advantages", with regard to science. The justification of formal logic is, at its foundation, the pragmatic consequence of such a language for science. This consequence is that the part of reality encapsulated in our natural language concerning propositions that changes truth-value must be left outside of formal logic as an aspect of how we experience reality; it cannot be part of any precise way of representing reality in logic:

> Earlier I suggested that Mr. Strawson's failure to appreciate the tenselessness of quantification over temporal entities might be a factor in his underestimation of the scope of modern logic. I should like to go further and say that I do not see how, failing to appreciate the tenselessness of quantification over temporal entities, one could reasonably take modern logic very seriously. (Quine 1953, p. 443, [14])

Quine and Strawson's discussion about the role of tenses in propositions sets the stage for the return of medieval logic in the philosophy of time. Formal logic, according to Strawson, cannot stand alone. The scope of logic is wider than what can be formalised and must also include what Strawson calls a "study of the logical features of ordinary speech." In such a study, we should, however, not expect to find "that character of elegance and system which belongs to the constructions of formal logic" (Strawson 2011, p. 232, [20]).

3 Prior's Invention of Tense-logic

Unknown to Quine and Strawson, two contemporary philosophers—J.J.C. Smart (1920–2012), also known as Jack Smart, from Australia and his friend A.N. Prior (1914–1969) from New Zealand—were engaged in a correspondence about the same topic in 1954. Prior and Smart had met each other in 1951, and a deep friendship formed between the two. Their correspondence left behind a series of letters now available at the Bodleian Library in Oxford (see [4]). In July 1954, Prior revealed to Smart his discovery of tense-logic for the first time. At the time, Prior was preparing a paper for the second New Zealand Philosophical Congress in Wellington on the discussion of the Master Argument, which was later published as (1955, [7]).[1]

Prior's discussion of the ancient argument begins with a clear reference to the way in which propositions were perceived in ancient and medieval logic: propositions may be true at one time and false at another. Prior's aim was to analyse the Megaric logician Diororus' understanding of possibility and necessity as "that which either is or at some time will be true, and the necessary as that which both is and always will be true" (Prior 1955, p. 1, [7]). Prior proposes to extend the logical vocabulary so that it not only contains the traditional proposition variables such as p, q, and r but also contains a set of operators that form functions taking these propositions as values. For the future tense "will be", Prior uses F and for the past "has been", he uses P; thus, Fp means "It will be the case that p" and Pp means "It has been the case that p". Pp and Fp are weak operators corresponding to a quantification over some time in the past or in the future. From these, a strong operator over the future and the past can also be defined.

Prior's treatment of the Master Argument demonstrates that if something is now true about the future, then it will necessarily be true ($Fp \supset \Box Fp$). To reach this conclusion, other tensed propositions must be assumed to hold. One is the assumption that propositions about the past are now unpreventably true ($Pp \supset \Box Pp$). Another assumption is a complex tensed proposition stating that "when anything is the case, then it has always been the case that it will be the case" ($p \supset \Box HFp$).[2] Prior's analysis is, thus, clearly medieval in the sense described above

[1] Prior spells it "Diodoran", while Benson Mates spells it "Diodorean".

[2] See [7].

by Øhrstrøm and Hasle, because it assumes the same intimate connection between logic and language concerning future contingency as do Ockham and Scotus.

On 30 July 1954, Smart wrote back to Prior to address Prior's suggestion of tense-logic. His response reveals that he was aware of Quine and Strawson's discussion.

> I don't feel the problem of Diodorus as a live one. Why shouldn't we say that what has happened might not have happened? Of course if the universe is deterministic and if by "impossible" we don't mean "ruled out by the laws of nature" but (rather eccentrically) "ruled out by the laws of nature + initial conditions" then what happens is "necessary" and what does not happen is "impossible". But "past" and "future" don't come into the matter—They only date events with respect to the moment we are at present discussing the matter in. I don't believe in any metaphysical difference between past and future—in fact I believe the assertion of such a difference can be refuted. And here I have Quine on my side - cf. his article on Strawson in Mind. (Smart, 1954 [18])

Smart has clearly chosen to adopt Quine's idea that tenses are tacit assumptions of a context by which a proposition is dated in a tenseless manner. Smart and Prior, thus, found themselves on either side of a divide concerning whether the Master Argument is a problem or whether it is not. The problem is only real if the tensed statements are logically well-formed propositions, which they, in Smart and Quine's eyes, are not. Despite Smart's inability to see how the past and future have anything to do with determinism, his reference to Quine and Strawson's discussion is helpful. It is highly likely that Prior only learned of Quine's discussion with Strawson from this letter. Furthermore, it is reasonable to assume that both the reference itself and Smart's response to Prior were the motivation behind Prior's *The Syntax of Time Distinctions* (1958), [9]. Its publication date obscures the fact that it was written as Prior's presidential address at the New Zealand Philosophical Congress given on 27 August 1954, barely a month after Smart's response. In *The Syntax of Time Distinctions* [9], Prior jumps right into Quine and Strawson's discussion, beginning his article with what he is aware will be the crucial assumption that sets his view on logic apart from that of both Strawson

and Quine: "Truth, on the face of it, is a property of propositions which is liable to alter with the time at which they are put forward" (Prior 1958, p. 105, [9]).

In this address, Prior first explains how this view on propositions not only was fundamental to medieval, as well as ancient logic, but sets these apart from modern logic. He then emphasises how there are no "grounds of a purely logical character for the current preference, and [...] 'propositions' in the ancient and medieval sense lend themselves as readily to the application of contemporary logical techniques and procedures as do 'propositions' in the modern sense" (Prior 1958, p. 105, [9]). Prior therefore sets out for himself the task of showing how the ancient and medieval conceptions of logic render both Quine and Strawson wrong.

Strawson is wrong, according to Prior, because he "regards it as a limitation of modern methods that they cannot cope with 'propositions' in the modern sense." (Prior 1958, p. 151, [9]). Prior intends to show that this is false, and that contrary to Strawson's view, one can by modelling the tenses in the abovementioned manner reach a conclusion as strong as those in modern modal logic. Quine, in contrast, is wrong when he "objects to the use of such 'propositions' in logic because modern methods cannot handle them" (Prior 1958, p. 105, [9]). It is not difficult to present Strawson's example in Prior's notation

(3) $P\exists x : (x \text{ is a woman} \wedge x \text{ is among the survivors})$.

Prior's past operator, P, with the meaning "it has been the case", allows him to use only the present-tensed "is" under the scope of the existential quantifier $\exists x$. The range of the quantifier is thereby no longer the entire space-time eternal continuum. It is rather the state of affairs that was the case when the woman survived. Formally, P is thus a propositional operator that logically takes us back to the time in the past when the statement

(4) $\exists x : (x \text{ is a woman} \wedge x \text{ is among the survivors})$.

was true.

It is obvious how the same can be done for the future, for which Prior introduced the operator F with the tensed meaning "it will be the case". Clearly, Prior's tense-logic, based on the ancient and medieval view of logic, is capable of providing a formal notation that renders a

formal analysis of tensed propositions much closer to natural language than tenseless language can.

4 The Return of Medieval Logic in the Philosophy of Time

There are four ways in which Prior's invention of tense-logic constitutes a return of medieval logic in the philosophy of time. The first is the obvious way outlined above in Prior's discussion of Strawson and Quine's argument: tense-logic constitutes a fundamental return to medieval logic by treating propositions in the same way that ancient and medieval logicians did, as logically well formed. Prior's development of tense-logic in *Time and Modality* (1957), [8] and *Past, Present and Future* (1969), [10], and the fact that temporal logic has proven highly valuable as a computational tool, cements Prior's point against Quine. Contrary to Quine, it is indeed possible to take modern logic seriously as tensed. Second, by returning to the way ancient and medieval logic treated propositions, Prior's work constitutes a correction of the mistaken view of medieval and ancient logic that was propounded through the work of Russell, who presented formal logic as discontinuous with the work of ancient and medieval logicians. Modern symbolic logic stands, in Prior's eyes, as a continuation of medieval logic, as his comment on Leibniz demonstrates:

> And of the founders of new philosophical schools, Leibniz, unlike most, was really interested in our subject, had considerable respect for the achievements in it of Aristotle and the Schoolmen, and anticipated modern developments by trying to find a symbolism which would give it a precise mathematical form, and make vast extensions of it possible.
>
> (Prior 1951, p. 17, [6])

In Prior's eyes, neither Russell nor Whitehead "brought to their work on mathematical logic any very serious close or detailed acquaintance with the logic of Aristotle and the Schoolmen" (Prior 1951, p. 17, [6]). This lack of knowledge, coupled with what Prior conceived as an attitude of contempt had, in his eyes, "helped to make customary in English-speaking countries for modern mathematical logic and the Aristotelian

logical tradition to be set in sharp opposition to one another" (Prior 1951, p. 46, [6]). Third, Prior's invention and development of tense-logic constitutes a fundamental challenge to Russell's perspective on analysis, which like medieval logic openly invites metaphysical questions back to the table. The question Strawson really poses is, "can a Russellian view on philosophical analysis hold up with regard to tenses?" His answer is that no, such an analysis cannot make sense of tenses. That conclusion is agreed upon by Quine, but Quine disagrees about whether that constitutes a challenge to a Russellian perspective on formal logic. The question is, however, on what basis Quine disagrees and says that one cannot claim to take modern logic seriously if one does not take it to be tenseless. Is it because a philosophical analysis of tenses *formally* requires a tenseless understanding of time, or is it because an analysis *metaphysically* requires the tenseless view on time? It cannot be the latter without contradicting Russell's view on analysis and, in light of tense-logic, it cannot be the former. Indeed, Prior's development of a logic for the tenses, and a parallel logic for tenseless dates, openly invites the metaphysical question of what is, metaphysically, most fundamental? It is not logic that decides this, but rather, to use a phrase by Prior, a "choice of the soul". (Prior 2003, p. 284 [13]) Consequently, it is legitimate within the philosophy of time to follow the medieval approach of analysing the importance of Aristotle's logic through questions about God's perspective on time. Our logical discussions aim at helping us decide on what we fundamentally take to be the truth, in a broad sense, and this includes our fundamental metaphysical commitments. Prior's own commitment to tenses was thus central to his logical work, and comes across like a theological creed:

> So far, then, as I have anything that you could call a philosophical creed, its first article is this: I believe in the reality of the distinction between past, present, and future. I believe that what we see as a progress of events is a progress of events, a coming to pass of one thing after another, and not just a timeless tapestry with everything stuck there for good and all. (Prior 1996, pp. 47-51, [11])

Fourth, and finally, tense-logic, by insisting on the formal validity of fundamental metaphysical commitments, openly invites metaphysical questions related to taking tenses seriously. Of these, the predominant

question is the one Prior worked on from 1953 to his death in 1969, the question of future contingency (see [5]). Smart's abovementioned response to Prior's disclosure of tense-logic reveals that Smart's failure to take tenses seriously was instrumental in his inability to perceive the Master Argument as constituting a genuine philosophical problem. For Smart "past" and "future" do not come into the matter—they only date events with respect to "the moment we are at present the matter in" (Smart 1954, p. 2, [18]). The converse, however, also holds: if "past" and "future" can be made formally acceptable—then ones position on metaphysical problems pertaining to time, like the problem of future contingency, will be legitimate objects of philosophical analysis. The importance of Prior's discovery of tense-logic and his application of it to the theological problem of God's foreknowledge and free will have been recognised by philosophers working with the problem of divine foreknowledge and human freedom, such as Hasker (1998), [3]. It was, indeed, Prior who, already in 1962 with *Formalities of Omniscience*, [11], began the modern discussion of the topic that occupied Scotus, Ockham and Thomas Aquinas in late medieval philosophy. Prior's correspondence with Smart provides a valuable window into the critical phase of Prior's discovery of tense-logic from when he first disclosed it to Smart in July 1954 until he presented tense-logic at the John Locke Lectures. In 1955, Prior was invited to give the lectures in 1956 at Oxford. At this point, in June 1955, he disclosed his plan to present tense-logic at the event to Smart, whose negative response demonstrates his belief that if Prior presented tense-logic, he would be running directly into a conflict. Thus, Smart warned Prior against it quite frankly:

> I still get the feeling that these lectures would more clearly represent your genius if you cut down on the metaphysics and stepped up the logic. As far as I can see you are at present trying to formulate ordinary tense-logic [...] This produces a pretty cumbersome system. That is, not a pretty system! (The tenseless logic is far superior aesthetically) [...] I would, honestly, strongly suggest cutting down on the quasi-metaphysics and increase the amount of formal logic.
>
> (Smart 1955, [18])

We find similar warnings repeated in the letters leading up to the John Locke Lectures, and Smart's influence on the development of the lec-

tures also comes through in *Time and Modality*(1957, [8]). They were warnings from a close friend who was honestly concerned that Prior's John Locke Lectures would not be the success they could be if Prior insisted on presenting tense-logic. Smart, following Quine, was not convinced that it could be characterised as formal logic, as is evident from his above warning. It was, as Smart perceived it, quasi-metaphysics.

Nonetheless, contrary to Smart's warnings, Prior went on to present tense-logic at the lectures, and it turned out to be a huge success. It launched Prior into his fame as perhaps one of the most important philosophers of that time in the late twentieth century. Part of his success could be a result of knowing that Oxford, the home of William of Ockham, would love a philosopher who did philosophical analysis in the same sense as he did. That tense-logic ultimately proved itself valuable in computer science, in terms of system checking and program verification, provides an interesting twist in the story of the influence of medieval logic on modern discussions in the philosophy of time. In *Past, Present and Future* (1967, [10]), Prior in a comment on tense-logic for discrete time noted that "the usefulness of systems of this sort does not depend on any serious metaphysical assumption that time is discrete; they are applicable in limited fields of discourse in which we are concerned only with what happens next in a sequence of discrete states, e.g. in the working of a digital computer." (Prior 1967, p. 67, [10]). In 1977 Prior's vision for the applicability of tense-logic to computer science was realized when the Israeli scientist Amir Pnueli applied Prior's concepts to program verification (see [22]). This development serves to underline the rightness of the intuition that kept Prior determined to bring tense-logic to Oxford, despite its obvious challenge to the then-prevalent view of philosophical analysis. It is, as Copeland writes, "pleasant to reflect that two major forces in the genesis of these software technologies were Prior's love of ancient and medieval logic and his concern to make conceptual room for freedom of the human will" (Copeland 2017, [1]).

Bibliography

[1] Copeland, B.J. (2017). "Arthur Prior". In: E.N. Zalta (ed.): *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition). https://plato.stanford.edu/archives/sum2017/entries/prior/.

Accessed 14 June 2018.
[2] Hintikka J.J. and Spade, P.V. (2017). "History of Logic". In: *Encyclopædia Britannica, Encyclopædia Britannica Inc*. https://www.britannica.com/topic/history-of-logic/Medieval-logic. Accessed 22 March 2018.
[3] Hasker, W. (1998). *God, Time and Knowledge*. Ithaca: Cornell Publishing.
[4] Jakobsen, D. (2017a). "The Significance of the Prior-Smart Correspondence for the Rise of Tense-Logic". In: Hasle, P., Blackburn, P. and Øhrstrøm, P.(eds.): *Logic and Philosophy of Time: Themes from Prior*. Aalborg University Press: pp. , 63–82.
[5] Jakobsen, D. and Øhrstrøm, P. and Hasle, P. (2017b). "In Celebration of Past, Present and Future". In: Hasle, P., Blackburn, P. and Øhrstrøm, P.(eds.): *Logic and Philosophy of Time: Themes from Prior*. Aalborg University Press: pp., 9–27.
[6] Prior, A.N. (1951). "The Craft of Logic". In: *The Prior Collection*, *Bodleian Library*, Box 22, Oxford 1951.
[7] Prior, A.N. (1955). "Diodoran Modalities". *The Philosophical Quarterly*: 205–13.
[8] Prior, A.N. (1957). *Time and Modality*. Oxford: Oxford University Press.
[9] Prior, A.N. (1958). "The Syntax of Time Distinctions". *Franciscan Studies* **18**(2): 105–20.
[10] Prior, A.N. (1967). *Past, Present and Future*. Oxford: Oxford University Press.
[11] Prior, A.N. (1996). "Some Free Thinking About Time". In: Copeland, B.J. (ed.): *Logic and Reality: Essays on the Legacy of Arthur Prior*. Oxford: Oxford University Press: pp. 47–51.
[12] Prior, A.N. (2003a). "The Formalities of Omniscience. In: Hasle, P., Øhrstrøm, P., Braüner, T. and Copeland, J. (eds.): *Papers on Time and Tense*. Oxford 2003: pp. 39–59.
[13] Prior, A.N. (2003b). "Modal Logic and the Logic of Applicability". In: Hasle, P., Øhrstrøm, P., Braüner, T. and Copeland, J. (eds.): *Papers on Time and Tense*. Oxford 2003: pp. 275–92.
[14] Quine, W.V. (1953). "Mr. Strawson on Logical Theory". *Mind* **62**(248): 433–51.
[15] Russell, B. (1914). *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*. Chicago, The open court publishing

company.
[16] Russell, B. (1959). *My Philosophical Development*. London: George Allen & Unwin.
[17] Russell, B. (1992). *A Critical Exposition of the Philosophy of Leibniz*, 2nd edn. 1937; repr. London: Routledge, 1992, (1900).
[18] Smart J.J.C. (1954). "Letter to A.N. Prior", 30th July, 1954. *In The Prior Collection*, *Bodleian Library*, Box 3, Oxford.
[19] Smart J.J.C. (1955). "Letter to A.N. Prior", 19th June, 1955. *In The Prior Collection*, *Bodleian Library*, Box 3, Oxford.
[20] Strawson, P.F. (1952). *Introduction to Logical Theory*, New York, 2011.
[21] Uckelman, S. (2009). *Modalities in Medieval Logic*, Institute for Logic, Language and Computation, Amsterdam.
[22] Øhrstrøm, P. and Hasle, P. (1995). *Temporal Logic from Ancient Ideas to Artificial Intelligence*, Dordrecht.

The Significance of the Contributions of A.N. Prior and Jerzy Łoś in the Early History of Modern Temporal Logic

Peter Øhrstrøm
Aalborg University, Denmark
poe@hum.aau.dk

Per Hasle
University of Copenhagen, Denmark
per.hasle@hum.ku.dk

Abstract

A.N. Prior is commonly recognized as the founding father of modern temporal logic. However, in a recent paper Marcin Tkaczyk and Tomasz Jarmużek have argued that temporal logic was in fact invented by Jerzy Łoś and first presented in 1947. In the present paper it is pointed out that the evaluation of this claim depends on the definition of the notion of temporal logic, and that there is a straightforward and commonly accepted understanding of temporal logic, according to which Prior is unquestionably the founding father of modern temporal logic. Furthermore, Tkaczyk and Jarmużek have maintained that "Prior was aware of and inspired by Łoś's ideas when beginning his own works in the field". It is unknown how much (if anything at all) Prior knew about Łoś's ideas before August 1954, when he presented his first version of temporal logic at a conference in Wellington, but nothing suggests that Łoś's ideas had any significant influence upon this original presentation. However, in his later writings Prior took Łoś's contributions into serious consideration. Prior appreciated several of Łoś's findings, but he wanted to defend an approach to

the field that went much further that the approach suggested by Łoś.

Keywords: Temporal logic, tense-logic, A.N. Prior, Jerzy Łoś.

1 Introduction

A.N. Prior is commonly regarded as the founding father of modern temporal logic. However, Marcin Tkaczyk and Tomasz Jarmużek (2018, [13]) have recently argued that temporal logic was in fact invented and indeed founded by Jerzy Łoś (1920–1998), since he "constructed, described and examined the first mature calculus of temporal logic" (p. 1, [13]). Furthermore, Tkaczyk and Jarmużek have maintained that "Prior was aware of and inspired by Łoś's ideas when beginning his own works in the field". As sources of the claim that Prior is the founding father of temporal logic, the authors mention a number of publications by Øhrstrøm and Hasle (and only those), but it should be pointed out that they share this role with many others. For instance, in their authoritative article on 'Temporal Logic' in the *Stanford Encyclopedia of Philosophy*, Goranko and Galton [1] briefly discuss Prior's inspiration from classical and medieval discussions on determinism, foreknowledge, etc., and go on to observe that

> From the early 1950s Arthur Prior set out to analyse and formalise such arguments, leading him, inter alia, to the invention of formal Temporal Logic, several versions of which are discussed below. Prior's seminal work, partly influenced by important precursors such as J. Findlay, H. Reichenbach and J. Łukasiewicz, initiated the modern era of temporal logical reasoning. (Goranko and Galton 2015, [1])

Before critically examining the contentions of Tkaczyk and Jarmużek (2018, [13]), it should be acknowledged that their paper is a valuable reminder about Łoś and his importance, and especially his early ideas about time and logic. It should certainly be conceded that it would have been an improvement to works such as (Øhrstrøm and Hasle 1993, 1995, 2006a, 2006b, [14-17]), as well as other works involving the history of temporal logic, if they had included reference to Jerzy Łoś. It will for future works with this subject matter be proper and worthwhile to point

out that Łoś (1947, [4]) contained a detailed and important presentation of the "Utp notation" as Prior calls it in his *Time and Modality* (1957, p. 20, [9]]) dealing with statements like Utp (p is satisfied in t). There is no doubt that this work should be seen as a remarkable contribution in the very early history of modern temporal logic. It is also true – as made quite clear several places in Prior (1957, [9]) - that Prior by then was acquainted with Łoś (1947, [4]) and considered these ideas. Certainly Prior did not attempt to hide the fact that at least by 1955, he had learned about the Utp notation suggested by Łoś. Although an English version of the 1947 paper had to wait until 1977 [5], H. Hiż published an English review of Łoś's paper in 1951 [2]. However, to concede these points is very far from conceding that Łoś could or should be regarded as the founding father of temporal logic. First and foremost, we have to clarify two issues: the very notion of temporal logic, and the degree of influence Łoś could have had on Prior's development of temporal logic. In Section 2 it will be argued that according to a straightforward and commonly accepted understanding of "temporal logic", Prior should clearly be regarded as the inventor of modern temporal logic. In section 3 it will be pointed out that in Prior's very first presentation of temporal logic in August 1954 Łoś's work was not mentioned. It is unknown whether he at that time was acquainted with Hiż's brief review from 1951, but nothing indicates that Łoś's work has had any significant influence upon Prior's first presentation of temporal logic. However, later on he clearly took Łoś's work into serious consideration. He appreciated many of Łoś's ideas, but he felt that he would rather put them into a much broader temporal context.

2 The Notion of Temporal Logic

According to Goranko and Galton (2015, [1]) the term "temporal logic" is sometimes simply identified with "tense-logic", whereas the term understood in a broader sense covers "all approaches to representation and reasoning about time and temporal information within a logical framework". In other words, in this broader sense temporal logic should offer a formal language rich enough to deal with all philosophical and scientific discussions that take the temporal aspects of reality into account. Following Prior's historical analysis (1957, p. 104 ff, [9]) this kind

of logic may be conceived as a re-invention of the view of logic in Antiquity and the Middle Ages, now with the added power (but also new challenges) of symbolic logic. According to this view the central tenets of Medieval logic with respect to time and tense can be summarised in the following way:

(1) tense distinctions are a proper subject of logical reflection, and
(2) what is true at one time is in many cases false at another time, and vice versa.

Understood in this manner a logical framework should fulfill both (1) and (2) in order to qualify as temporal logic. Among other things, temporal logic should offer a conceptual framework for the famous discussions regarding McTaggart's A- and B-series and their mutual relations. It is worth noting that (1) and (2) run counter to well-established post-medieval ideas: namely (a) that logic should be developed as a timeless science, and (b) that propositions are really incomplete without an explicit time-reference (thus "Socrates is sitting" would be regarded as incomplete until a time-reference was supplied) (cf. Øhrstrøm and Hasle 1995, pp. 109–117, [15]). It was part of the goal of Prior's work to overturn these well-entrenched positions (i.e. (a) and (b)).

In Hiż' review, Łoś's basic idea was presented in the following manner: "The main purpose of this paper is to analyse Mill's canons as rules of operation for a part of the language of physics. To do it the author builds up an axiomatization of a fragment of the physical language." (Hiż 1951, p. 58, [2]) The suggestion seems to be that the *Utp* notation can serve as a precise representation of the reasoning about time as needed in physics. The paper by Łoś is certainly interesting and important. And the *Utp* notation is, of course, an important element in a fullfledged temporal logic. However, there is no genuine analysis of tenses in Łoś's paper – probably because tenses do not appear to be needed in physics. Marcin Tkaczyk and Tomasz Jarmużek state: "The only formal tool Łoś's work [1947] lacks of is representing tenses by means of modal connectives. This idea Prior took from Findlay." (Tkaczyk and Jarmużek 2018, p. 16 [13) This means that although Łoś's 1947 paper fulfills a version of (2) above, it does not live up to (1). As indicated by Tomasz Jarmużek and Andrezej Pietruszczak (2004, [3]), what Jerzy Łoś did in the field was that he "proposed an operator that referred sentences to temporal moments" (2004, p. 147, [3]).

The lack of a proper representation of tenses is in fact crucial. However, the re-introduction of tenses into logic is an utterly essential feature of Prior's work. It may not be needed in physics as we know it, but Prior wanted to go beyond physics and he insisted that tenses are needed in order to deal with the notion of existence and indeed with the very understanding of time. In passing, it should be noted that the statement 'This idea Prior took from Findlay' is an exaggeration, Prior's generous references to Findlay notwithstanding (cf. Øhrstrøm and Hasle, 1995, pp. 170–71, [14]).

Prior wanted to develop a logic based on tenses, p (it was the case that...) and F (it will be the case that (or in the metricized version, Pn (it was the case n time units ago that ...) and Fn). As quite explicitly stated in Prior 1957 [9], it was in this connection also his ambition to show that the content of Łoś's Utp notation can be formulated in terms of a rather simple system of tense-logical postulates (p. 20, [9]). In other words, what Prior wanted was to formulate ideas that could supplement the approach to temporal reasoning normally used in physics.

3 Prior and Łoś

Prior first presented his temporal logic at a conference in Wellington in August 1954. This very first presentation of his ideas on temporal logic was later published in Prior (1958, [10]). We don't know exactly when Prior became aware of Łoś's work, but it was briefly mentioned in the last section of Appendix 1 in Prior's *Formal Logic*, where Łoś's ideas are introduced under the headline "Logic of Assertion and Formalized Physics" (Prior 1955, p. 313, [8]). Furthermore, there is a reference to the review by Hiż in Prior's *Time and Modality* (1957, p. 20, [9]). As pointed out by Tkaczyk and Jarmużek (2018, [13]), Hiż did a brilliant condensation of Łoś's work in his review, but even so it was very brief. In *Past, Present and Future*, Prior stated that "I only know the paper [i.e. Łoś' 1947, [4]] through Hiż' review" (cf. Prior, 1967, p. 212, [11]), and in any case Łoś' paper was as already noted not available in English before 1977. Prior's 1954 presentation does not contain any reference either to Łoś or to Hiż. The presentation contained a discussion of a logic similar to the Utp notation suggested by Łoś, but the term used by Prior is the l-calculus (referring to "later") and furthermore he used 'px' to stand for 'p at x'. If Prior had in fact read and been influenced by the

review of Łoś's system before August 1954 as suggested by Tkaczyk and Jarmużek, he would by any reasonable assumption have made a reference to Łoś in the paper. For one thing, he treasured Polish logic very much and promoted it throughout his life; for another thing, it would have been advantageous to refer to this Utp notation instead of making up his own "px notation". Indeed, had Prior been acquainted with Hiż' review' at that time, he would have had every reason to use this reference. Apart from the fact that Prior was conscientious, sometimes almost to excess, about crediting his influences, he needed all the powerful references he could draw on when making his case for a formal logic of time. This idea was in 1954 not only unorthodox but also somewhat contentious.

From the historical perspective, it is also worth pointing out that Prior had entertained ideas about time and logic as early as 1950/51. In 1951, he submitted the mammoth manuscript *The Craft of Formal Logic* to Oxford University Press. (The manuscript was never published in its 1951-form, but was replaced by Prior's *Formal Logic* [8]). In her recent investigation, Aneta Markoska-Cubrinowska observes:

> Arthur Prior's unpublished manuscript "The Craft of Formal Logic", written in 1950–51, contains his early ideas about time as a semantic concept. Years before he would publish his tense logic, Prior contemplated the construction of a semantic theory of propositional logic in which propositions are interpreted as functions of time instants. These ideas are born from reviewing historical material, and particularly from his analysis of Diodorus's conditional and Boole's propositional algebra. He suggests that 'P entails Q' could be expressed formally as '$\forall i(Pi \to Qi)$', where 'i' stands for an instant of time, and 'Pi' stands for 'p is true at i'.
>
> (Markoska-Cubrinowska 2017, p. 29 [6])

In the paper's conclusion, Aneta Markoska-Cubrinowska sums up the paper's results as follows:

> Prior's unpublished manuscript "The Craft of Formal Logic", written in 1950–51, reveals that his development of temporal semantics had begun at least three years before he first publicised his tense logic.
>
> (Markoska-Cubrinowska 2017, p. 39 [6])

Apart from the considerable historical interest that this investigation carries, it also implies another striking point. Prior himself dated the beginning of his development of tense-logic, respectively, temporal logic, to 1953, when he came across Findlay's footnote (cf. Øhrstrøm and Hasle 1995, pp. 170–71, [14]). Thus he did not even see his own earlier ideas about time and logic – involving instants, quantification over instants and the notion 'p is true at i' – as a proper anticipation of his development of tense-logic from 1953 and onwards.

In his *Time and Modality*, Prior clearly recognized Łoś as the first scholar to develop the Utp notation. Furthermore, in this book Prior presented a detailed analysis of the axiomatic account of the Utp notation as it was suggested by Łoś (Prior 1957, pp. 19-21, [9]). In his discussion of the conceptual and philosophical aspects of Łoś's logic Prior wrote: "Even Łoś's logic is not, indeed, a tense-logic but rather an unanalysed date-and-interval logic, but it is at least a logic in which the time-reference is made by an operator which takes whole 'predications' as its arguments." (Prior 1957, p. 107, [9]) Prior also noted that Łoś's date-and-interval logic may be seen in the light of his earlier discussion of a 'logic of assertion', i.e. the logic of statements of the form 'x asserts that p'. Prior appreciated Łoś's findings and stated: "[...] he has found, as no one before him seems to have found, an appropriate symbolism for this type of subject-matter, and he sees that the logic of dates and intervals and the logic of assertion both require a symbolism of this kind." (Prior 1957, p. 122, [9]).

In Chapter 1 of his *Past, Present and Future*, Prior (1967, [11]) names and discusses important precursors of modern tense-logic. Unfortunately, Łoś' contribution is not included in this chapter. However, in Appendix B of *Past, Present and Future*, Prior himself characterizes this omission as a mistake. As a kind of compensation, Prior includes a whole section in the appendix dedicated to the works of Łoś (1967, pp. 212-213, [11]). Here Łoś is not only acknowledged as a precursor of modern tense-logic, but also as a precursor of the logic of instants – and in fact also of what we now call hybrid logic. In particular, Prior emphasizes the importance of the so-called clock axiom, which was in fact mentioned already in the brief section in Formal Logic (Prior 1955, p. 313, [8]). According to this axiom there is a proposition p for any instant a, so that for any instant b, the proposition p holds at b if and only if a is identical with b, i.e.

$$\exists p \forall b : Tbp \equiv (a = b)$$

Prior states: "Łoś regarded it as 'our only weapon against the metaphysical and extrasensual conception of time'. His point of view seems in fact to have been very close to that of Sections 3 and 4 of this appendix. The clock axiom, one might say, might justify (or might reflect) our *identification* of an 'instant' with a proposition true at that instant only [...]" (Prior 1967, p. 213, [11]).

4 Conclusion

The paper by Tkaczyk and Jarmużek is a welcome reminder that there was a well developed anticipation of a highly important part of temporal logic, the Utp notation, available in 1947 – of which Jerzy Łoś was the author. This reminder could and should enrich future expositions of the development of temporal logic, and we can only be thankful for this contribution. However, the paper definitely fails in establishing that Jerzy Łoś should be regarded as the founding father of temporal logic. The preceding conceptual and historical analyses are certainly the most important arguments in this matter, but one may also stop to ponder for a moment what it rightfully means to be 'founding father'. The history of science is full of examples of great intellectuals arriving independently at similar or even identical ideas at various times. However to be recognized as founder one also has to be the one who develops these ideas and establishes them in the scientific community. One has to hold a crucial role in what is called in German the "Wirkungsgeschichte", i.e. the history or process of influence. Jerzy Łoś regrettably did not follow up his brilliant version of the Utp notation, but left the idea in favour of other intellectual pursuits. There is no evidence that he exerted any influence on the development of temporal logic before Prior's *Time and Modality*, and as already stated this influence was mainly confined to the discussions of the Utp notation. One could compare with the contribution of Hans Reichenbach, also from 1947. Reichenbach's work also contained brilliant ideas on time and logic and has had considerable influence in later discussions. While Reichenbach's proposals were far less developed than Łoś in terms of formal maturity, he however had a keener eye than Łoś when it came to the finer details of temporal distinctions. But noone has ever argued that Reichenbach, in this respect Łoś'

exact contemporary, should be seen as the founding father of temporal logic. He was, as was Łoś, a precursor.

Seeing modern temporal logic as a re-invention of the logic in Antiquity and the Middle Ages but now developed in terms of symbolic logic, Prior's account of the history of tense-logic suggests that the rise of temporal logic would take place in two steps corresponding to the steps that characterized the downfall of medieval tense-logic (cf. Øhrstrøm & Hasle 1995, p. 109, 15]). However, in the context of the rise of modern temporal logic one should expect that the two steps would appear in the reverse order of the order discernible during the downfall of medieval tense-logic, i.e. the order of the resurgence should be expected to be

(2) what is true at one time is in many cases false at another time, and vice versa.
(1) tense distinctions are a proper subject of logical reflection.

As noted earlier, in order to discuss all aspect of temporal reasoning (including important conceptual distinctions on time as for example those found in McTaggart's paradox), both (2) and (1) are needed. It is evident that the logic of dates and intervals suggested by Jerzy Łoś in 1947 fulfills (2), but not (1). In this way Łoś's logic may qualify as an important anticipation of the re-introduction of temporal logic. On the other hand, we have no evidence that A.N. Prior was aware of this when he in August 1954 for the first time presented a modern and symbolic logic fulfilling both (2) and (1).

Despite the qualities of the achievements of Łoś in 1947 as a precursor of tense-logic and even hybrid logic, it is obvious that Prior retains the honour of being the founding father of modern temporal logic.

Bibliography

[1] Goranko, V. and Galton, A. (2015). "Temporal Logic". In: Zalta, E.N. (ed.): *The Stanford Encyclopedia of Philosophy* (Winter 2015 Edition). https://plato.stanford.edu/archives/win2015/entries/logic-temporal/. Accessed Nov 28 2018.

[2] Hiż, H. (1951). Rev. J. Łoś, Podstawy analizy metodologicznej kanonow Milla. *Journal of Symbolic Logic* **16**: 58–59.

[3] Jarmużek, T. and Pietruszczak, A. (2004). Completeness of minimal positional calculus. *Logic and Logical Philosophy* **13**: 147–162.

[4] Łoś, J. (1947). Podstawy analizy metodologicznej kanonow Milla. *Annales Universitatis Mariae Curie* Sectio F, vol. II, 5: 269–301.
[5] Łoś, J. (1977). The foundations of a methodological analysis of Mill's methods. In: Przełęcki, M. and Wojcicki, R. (eds.): *Twenty-Five Years of Logical Methodology in Poland*. D. Reidel Publishing Company: pp. 291–325.
[6] Markoska-Cubrinovska, A. (2017).Time and Truth in "The Craft of Formal Logic": Prior's early temporal semantics born from reflexions on Diodorus and Boole. In: Hasle, P., Blackburn, P. and Øhrstrøm, P. (eds.): 2017. *Logic and Philosophy of Time: Themes from Prior*. Aalborg: Aalborg University Press: pp. 29–39.
[7] Prior, A.N. (1950-51). *The Craft of Formal Logic*. Unpublished manuscript: Kept in the Bodleian Library (Weston), Prior's Papers, Box 22.
[8] Prior, A.N. (1955). *Formal Logic*. Oxford: Clarendon Press.
[9] Prior, A.N. (1957). *Time and Modality*. Oxford: Oxford University Press.
[10] Prior, A.N. (1958). The Syntax of Time-Distinctions. *Franciscan studies* **18**: 105–120.
[11] Prior, A.N. (1967). *Past, Present and Future*. Oxford: Oxford University Press.
[12] Reichenbach, H. (1947). Elements of Symbolic Logic. *Journal of Philosophy* **45**(6): 161–166.
[13] Tkaczyk, M. and Jarmużek, T. (2018). Jerzy Łoś positional calculus and the origin of temporal logic. *Logic and Logical Philosophy*. DOI: 10.12775/LLP.2018.013. Published online: September 23, 2018.
[14] Øhrstrøm, P. and Hasle, P. (1993). A.N. Prior's rediscovery of tense logic. *Erkenntnis* **39**: 23–50.
[15] Øhrstrøm, P. and Hasle, P. (1995). *Temporal Logic: From Ancient Ideas to Artificial Intelligence*. Dordrecht: Kluwer Academic Publishers.
[16] Øhrstrøm, P. and Hasle, P. (2006a). A.N. Prior's logic. In: Gabbay, D.M., and Woods, J. (eds.): *Handbook of the History of Logic*, vol. 7. Amsterdam: Elsevier: pp. 399–446.
[18] Øhrstrøm, P. and Hasle, P. (2006b). Modern temporal logic: The philosophical background. In: Gabbay, D.M. and Woods, J. (eds.): *Handbook of the History of Logic*, vol. 7. Amsterdam: Elsevier: pp. 447–498.

B-theory and Time Biases

Sayid R. Bnefsi
University of California, Irvine
sbnefsi@uci.edu

Abstract

We care not only about what experiences we have, but when we have them too. However, on the B-theory of time, something's timing isn't an intrinsic way for that thing to be or become. Given B-theory, should we be rationally indifferent about the timing *per se* of an experience? In this paper, I argue that B-theorists can justify time-biased preferences for pains to be past rather than present and for pleasures to be present rather than past. In support of this argument, I appeal to the doctrine of temporal parts or "four-dimensionalism" for short. When held in conjunction with a certain evaluative principle about *whose* experiences matter, four-dimensionalism reconciles B-theory with some time-biased preferences.

Keywords: B-theory, time bias, four-dimensionalism.

1 Introduction

When were you last in great pain? That might not be worth remembering. For what it's worth, that's all in the past now—or so the saying goes. According to the B-theory of time, however, that might not be saying much. B-theory denies that something's timing is an intrinsic way for that thing to be or become. Instead, B-theory holds that *when* something happens is not metaphysically importantly different from *where*

it's happening. Yet in conceiving time and space as parallel metaphysical dimensions, must B-theory also conceive them as parallel evaluative dimensions?

All else equal, we should be and typically are rationally indifferent about a pain's whereabouts. Being in pain here, wherever I am, wouldn't be evaluatively importantly different from being in pain there, wherever you are. Call this *spatial neutrality* about experiences. Dissimilarly, some philosophers disagree that we should be rationally indifferent about a pain's timing. They believe that being in pain now, in the present, is importantly different from being in pain earlier, in the past.[1] Indeed, it seems most people prefer their pains to be past rather than present and their pleasures to be present rather than past. These preferences exemplify temporal value asymmetries, or "time biases" for short, which represent the various ways in which we care not only about what experiences we have, but when we have them too.

Many philosophers have claimed that B-theory undermines the rationale for being time biased and strengthens the rationale for *temporal neutrality*: roughly, the thesis that we should be rationally indifferent about an experience's timing *per se*.[2] The arguments from B-theory to temporal neutrality might have something like the following premises as their common ground. According to B-theory, time has an ontologically homogeneous structure, i.e., being past, present, and future are extrinsic ways for things to be or become. If something's timing is an extrinsic way for it to be or become, then we should care about when an experience happens *per se* to the same extent that we should care about where an experience happens *per se*. But we shouldn't care about where an experience happens *per se*. Therefore, according to B-theory, we shouldn't care about *when* an experience happens *per se*.

However, I will argue that B-theory does not necessarily predict that all time biases are irrational. This argument requires that B-theorists maintain a certain view about diachronic identity: namely, four-dimensionalism. According to four-dimensionalism, persistence through time is like extension through space. This means that facts about us over time are given by facts about our temporal parts at various times, in the same vein that facts about us at a time are given by facts about our spatial parts

[1] See Prior (1959, [16]), Parfit (1984, [13]), and Hare (2007, [7]; 2009, [8]).

[2] See Cockburn (1998, [2]), Zimmerman (2008, [20]), Greene and Sullivan (2015, [5]), and Pearson (2018a [13], 2018b [14]).

at that time. In the next section, I'll explicate an infamous objection against B-theory's compatibility with reasons for being time biased. In turn, I'll develop and outline my argument from four-dimensionalism for B-theory's compatibility with reasons for being time biased.

2 Thank Goodness That's Over

Some philosophers doubt that the correct semantic analysis for an area of inquiry indicates the correct ontological analysis for that area, but some B-theorists took this for granted in their arguments for their view.[3] These arguments appealed to certain reductive analyses of linguistic tense, but it is in effect an argument from parsimony. The argument's main premise is that the linguistic category of tense is available for paraphrase in entirely tenseless, but synonymous language. And that premise is supposed to support the conclusion that what it is for something to be past, present, or future just is for that thing to stand in certain tenseless, frame-relative relations: respectively, the earlier-than, simultaneous-with, or later-than relations. Whatever that parsimony argument's validity, the main premise has come under fire. And in that connection, most of the smoke has come from an objection made familiar by A. N. Prior:

> One says, e.g. "Thank goodness that's over!" ... [which] certainly doesn't mean the same as, e.g. "Thank goodness the date of the conclusion of that thing is Friday, June 15, 1954", even if it be said then. (Nor for that matter, does it mean "Thank goodness the conclusion of that thing is contemporaneous with this utterance". Why should anyone thank goodness for that? (Prior 1959, p. 17, [15]).

Initially considered, Prior's objection is that tenseless analyses of tense—for example, the date and token-reflexive analyses—fail to preserve the cognitive relations between tensed propositions and the propositional attitudes which embed them. Therefore, arguments in support of B-theory from the semantic eliminability of tense are unsound. But there's another objection against B-theory to which Prior gestures to-

[3] See Dyke (2007, [4]).

wards when, at the end of the foregoing passage, he raises the question: "Why should anyone thank goodness for that?"

When we fill in the anaphora, what Prior seems to be asking is this: "Why would a tenseless fact about the timing of an experience ever give someone a reason to value it differently?" In the next section, I'll offer a novel answer to the question that I take Prior to be raising against the B-theorist. Subsequently, I'll compare my answer to, and defend its superiority against, comparable answers due to D. H. Mellor [10-11], Murray MacBeath [9], and Heather Dyke and James Maclaurin [4].

3 Thank Goodness That's No Longer Me

According to four-dimensionalism, people are temporally extended wholes whose lifetimes are given by their moment-bound temporal parts. More specifically, four-dimensionalism maintains that there exist moment-bound things, i.e., temporal "parts" or "stages," just as there exist spatial parts of things. These temporal parts are what comprise temporally extended things, i.e., continuants, and facts about continuants supervene on facts about the temporal parts of the continuant.[4]

Four-dimensionalists can disagree about what continuants are, whether they have their properties derivatively, and the kind of relations that underlie the various facts about our properties over time. Worm theorists, or "perdurantists," maintain that the continuants over which we quantify and talk about are aggregates of moment-bound individuals or "worms" who derive their properties from the properties their temporal parts have. Stage theorists, or "exdurantists," agree that spacetime worms exist, but they identify the continuants over which we quantify and talk about with the individual stages of the worm, who have properties at other times by standing in particular counterpart relations to other stages of the worm at other times.[5] For the intents and purposes of reconciling B-theory and the rational permissibility of some kinds of time biases, not much weighs on the ontological and semantic differences between perdurantism and exdurantism. That said, I'll be raising my argument according to the perdurantist's account of diachronic

[4]See Balashov (2010, [1]) for a scientifically informed investigation into the debate about persistence.

[5]Balashov (2010, p. 13, [1]).

identity.

The first substantive premise of my argument is that B-theory in conjunction with four-dimensionalism entails that any experience that is over at a time for some temporal part also has the property of not belonging to that temporal part at that time. To illustrate, imagine that all of yesterday someone—Student—suffered from a mild headache, which is over today. According to B-theory, for Student's headache to be over just is for that headache to conclude simultaneously with respect to some frame of reference. And according to four-dimensionalism, for Student's headache to conclude today just is for that headache to belong to Student's temporal parts yesterday, but not to Student's temporal parts today. Student's temporal parts yesterday—Student-Yesterday—is not numerically identical to Student's temporal part today, Student-Today. Therefore, for Student's headache to be over at a time, if that is a tenseless feature of the headache, is also for the headache to belong to Student-Yesterday rather than Student-Today.

The second substantive premise of my argument involves a certain evaluative principle with respect to pleasures and pains inspired by Caspar Hare's [7, 8] work on biases towards our own favor or "self-bias" for short. The principle provides the normative reason for which one might "thank goodness" for the fact that some pain is past if that is merely a tenseless fact:

> **Mild Egocentric Hedonism** (**MEH**): All other things being equal, one should prefer a pain that is not one's own rather than one's own. Conversely, one should prefer a pleasure that is one's own rather than not one's own.

It is tempting, but *incorrect*, to confuse MEH above with this principle below:

> **Egocentric Schadenfreude** (**ES**)**:** All other things being equal, one should prefer a pain that is someone else's. Conversely, one should prefer a pleasure that is no one else's.

Here's the difference between the two. Imagine a group of restaurant workers drawing straws to decide who should clean the restrooms. Mild egocentric hedonists prefer not to draw the short straw themselves. Sure, by preferring not to draw the short straw, one prefers by extension that

someone else does. But it's the *schadenfreuder* who finds preference fulfillment not only in not drawing the short straw, but also in someone else drawing it. If they were able to choose, the mild egocentric hedonist can consistently prefer that no one, including them, have some negative experience. In contrast, the egocentric schadenfreuder is satisfied by the thought that someone experiences something unpleasant.

With MEH in tow, my argument for the reconciliation between B-theory and time biases continues as follows. For convenience, I put the argument in premise-conclusion form.

> **Premise 1, B-theory**: What it is for one's pain to be over is for its conclusion to be simultaneous with some current reference frame (e.g., an utterance).

In turn, we assume four-dimensionalism:

> **Premise 2, 4D**: Continuants are temporally extended beings who exist at various times with different qualities over time in virtue of having temporal parts of different qualities who exist at various times.

The sub-argument for the next premise was provided earlier.

> **Premise 3:** Any experience that is over at a time for some temporal part also has the property of not being experienced by that temporal part.

And the next premise provides the reason for which an experience that is over at a time is something for which one should thank goodness, given B-theory and four-dimensionalism.

> **Premise 4, MEH:** Whatever painful experience has the property of not belonging to some temporal part at a time would satisfy that temporal part's mild egocentric hedonism.

From the foregoing premises we have this to conclude:

> **Conclusion:** A continuant should thank goodness that some pain is over at a time, if that is a tenseless fact, because that continuant has a temporal part at some time whose mild egocentric hedonism is fulfilled in virtue of that tenseless fact obtaining.

Finally, it's worth illustrating my argument's logic. Consider *Dentistry*:

> Student is scheduled for a routine checkup with their dentist, Dentist. Of course, Student dislikes seeing the dentist. To provide them with some relief, Dentist says: "It'll all be over soon." As a B-theorist, however, Student puts Dentist's time bias into question: "And why should anyone thank goodness for that?" Dentist replies: "Because *then* it will no longer be *your* problem!"

4 Prospects for Temporal Relativity

Call my argument in the previous section the "Thank Goodness That's No Longer Me" argument. Not only does it answer Prior's question, but it also addresses a family of objections against the normativity of B-theory. Consider, for instance, the objection that David Cockburn raises:

> To show how anything like our current emotional life might be consistent with the claim that 'past', 'present', and 'future' events all have exactly the same kind of reality...His problem, that is, is to show how familiar ways in which we offer 'the fact' that something *has* happened, *is* happening, or *will* happen as a reason for actions and feelings can be acceptable if 'there are no tensed facts'.
>
> (Cockburn 1998, p. 85, [2])

More recently, Preston Greene and Meghan Sullivan imply that, for the purposes of rational evaluation, when an experience happens *per se* is not importantly different from where it is happening *per se*:

> Given the B-theory of time, distinguishing between past and future experiences can seem just as arbitrary as distinguishing between experiences that happen here and experiences that happen there.
>
> (Greene and Sullivan 2015, p. 953, [6])

The claims being made in the foregoing passages have a common denominator. Since B-theory implies that there is no intrinsic or absolute way to distinguish between past, present, and future events, it seems

that there is no reason to be responsive to an experience's timing for the purposes of rational evaluation. However, even if there is no intrinsic way to distinguish between past, present, and future events, nevertheless there may be some feature about an experience's pastness, presentness, or futurity—if those are tenseless ways for an experience to be—that's evaluatively relevant; that is what my Thank Goodness That's No Longer Me argument shows.

To show the unique contribution my argument makes, it is worth discussing how the argument interacts with similar arguments in the literature on the compatibility between B-theory and temporal value asymmetry. In response to Prior, among other A-theorists, Murray MacBeath (1994, [9]), with D. H. Mellor concurring (1993, [11]), presented an account of tensed propositional attitudes that distinguishes their formal objects from the content that constitutes them. On this account, the formal object of our relieving *belief* that a pain is over is a tenseless proposition or fact about that pain. But the object of one's *relief* in the belief that a pain is over is that belief's irreducibly tensed content. One thus has reason to feel relief in the belief that a pain is over rather than present or forthcoming not because that belief's formal object is an irreducibly tensed fact, but because of that belief's irreducibly tensed *content*.[6]

Heather Dyke and James Maclaurin (2002, [3]) have also presented an account of the compatibility between B-theory and temporal value asymmetry. They claim that our reasons for being time biased stem from considerations about natural selection, which themselves are ultimately responsive to the tenseless fact that the direction of causation is from earlier to later.[7] On their account, time-biased behaviors typically confer certain evolutionary advantages and fitness-enhancing effects due to causation's tenseless asymmetric direction. One thus has reason to feel relief in the belief that a pain is over not only because that belief's object or content is tensed, in contrast to Mellor and MacBeath, but also because relief is the evolutionarily useful attitude to have towards a past pain, and dread is the evolutionarily useful attitude to

[6]Let us illustrate the distinction. Suppose you are about to get into what looks like a fatal car accident. Of course, you scream in terror in the belief that you will die. Fortunately, you will not die. The object of your belief is a false proposition. But the object of your terror in the belief that you'll die is not the false proposition, rather it's that belief's terrifying content—the way the world would look like if the object of your belief were true.

[7]See Dyke and Maclaurin (2002, p. 285, [4]).

have towards a future pain.

Unlike my Thank Goodness That's No Longer Me argument, however, MacBeath and Mellor's argument, as well as Dyke and Maclaurin's, seem to confuse an important distinction between what we have *motivational* reason to do and what we have *normative* reason to do. That is, their arguments mistake what we are under psychological pressure to do with what we are under rational pressure to do. Indeed, reconsider MacBeath and Mellor's account. If the object of one's relief about past pains is some irreducibly tensed content, then the objects of relief, among other such tensed attitudes, are based on misrepresentations of reality if B-theory is true. But attitudes that are based on misrepresentations of reality are unjustified. Therefore, if B-theory is true, the various tensed attitudes that embody our temporal value asymmetries are unjustified.

In a similar vein, Dyke and Maclaurin's account also seems to imply that our tensed attitudes are unwarranted. Their account says that natural selection pressures explain why there are reasons to act in time-biased ways. But these natural selection pressures do not always track the truth. Being time biased may be a better way for a species to survive. That said, being a better way for a species to survive is not always a justified way for a species to behave. Unless there is an intrinsic connection between such natural selection pressures and the reasons that we have to evaluate experiences in certain ways rather than others, an evolutionary account of time bias seems to raise skepticism against the claim that being time biased is justified.

Finally, it is worth discussing a potential limitation of my view, but also foregrounding a potential solution. I have offered a tenseless account of some time biases. But as it stands, my account does not seem to justify our bias in favor of the future over the past with respect to pleasure and our bias in favor of the past over the future with respect to pains. Indeed, my argument only establishes that there are reasons to prefer present rather than non-present pleasures, and non-present rather than present pains. But both past and future pains are also non-present pains, so on my account distinguishing between them is arbitrary for the purposes of rational evaluation. Yet there seems to be something importantly different about a pain being future rather than past. To be sure, we are constantly under great psychological or evolutionary pressure to treat matters future and past differently. But being under

such pressure does not justify bias towards the future for the same reasons it does not justify bias towards the present.

In order to justify future bias in a manner compatible with B-theory, there must be something importantly different between the past and the future in virtue of which we have normative reason to prefer pleasures to be located in our future and pains to be located in our past. In this connection, there may be a parity argument according to which our temporal value asymmetries between the past and future are not importantly different from our personal value asymmetries between people with whom we are in variously intimate relationships. On this argument, just as it is not arbitrary to be more concerned about the experiences of certain relatives rather than others, i.e., one's nuclear rather than extended family, similarly it is not arbitrary to be more concerned about experiences that occur in certain parts of my lifetime rather than others, i.e., the future rather than past. In other words, in the interpersonal context, it seems our asymmetric attitudes between others are justified by the variously intimate relationships that we stand in with others, and the reasons that we have for standing in those relationships with them. Likewise, in the intertemporal context, it may be that our asymmetric attitudes between the past and the future are justified by the variously different metaphysical, epistemological, and ethical relations that we stand in with ourselves in the future rather than the past.

Bibliography

[1] Balashov, Y. (2010). *Persistence and Spacetime*. Oxford: Oxford University Press.

[2] Cockburn, D. (1998). *Tense and Emotion*. In: Le Poidevin, R. (ed.): *Questions of Time and Tense*. Oxford: Oxford University Press: pp. 77–91.

[3] Dougherty, T. (2015). Future-Bias and Practical Reason. *Philosophers' Imprint*, **15**(30): 1–16.

[4] Dyke, H. and Maclaurin, J. (2002). "Thank Goodness That's Over": The Evolutionary Story. *Ratio* **15**(3): 276–292.

[5] Dyke, H. (2007). *Metaphysics and the Representational Fallacy*. London: Routledge.

[6] Greene, P. and Sullivan, M. (2015). Against Time Bias. *Ethics*

125(4): 947–970.

[7] Hare, C. (2007). Self-Bias, Time-Bias, and the Metaphysics of Self and Time. *Journal of Philosophy* **104**(7): 350–73.

[8] Hare, C. (2009). *On Myself, and Other, Less Important Subjects*. Princeton: Princeton University Press.

[9] MacBeath, M. (1983). Mellor's Emeritus Headache. *Ratio* **25:** 81–88. Reprinted in: Oaklander, L.N. and Smith, Q. (eds.): *The New Theory of Time*. New Haven: Yale University Press, 1994: pp. 305–312.

[10] Mellor, D.H. (1981). "Thank Goodness That's Over". *Ratio* **23:** 20–30. Reprinted in: Oaklander, L.N. and Smith, Q. (eds.): *The New Theory of Time*. New Haven: Yale University Press, 1994: pp. 293–305.

[11] Mellor, D. H. (1983). MacBeath's Soluble Aspirin. *Ratio* **25**: 89–92. Reprinted in: Oaklander, L.N. and Smith, Q. (eds.): *The New Theory of Time*. New Haven: Yale University Press, 1994: pp. 312–316.

[12] Oaklander, L.N. and Smith, Q. (1994). *The New Theory of Time*. New Haven: Yale University Press.

[13] Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.

[14] Pearson, O. (2018). Appropriate Emotions and the Metaphysics of Time. *Philosophical Studies* **175**(8): 1945–1961.

[15] Pearson, O. (2018). Tensed Emotions, Evolution, and Time. *Philosophia* **46**(2): 401–409.

[16] Prior, A.N. (1959). Thank Goodness That's Over. *Philosophy* **34**(128): 12–17.

[17] Sider, T. (1997). Four-Dimensionalism. *The Philosophical Review* **106**(2): 197–231.

[18] Sider, T. (2001). *Four-Dimensionalism: An Ontology of Persistence and Time*. Oxford: Oxford University Press.

[19] Suhler, C. and Callender, C. (2012). Thank Goodness That Argument is Over: Explaining the Temporal Value Asymmetry. *Philosopher's Imprint* **12**(15): 1–16.

[20] Zimmerman, D. (2008). The Privileged Present: Defending an "A-theory" of Time. In: Sider, T., Hawthorne, J. and Zimmerman, D. (eds.): *Contemporary Debates in Metaphysics*. Blackwell: pp. 211–225.

Presentism and Cross-Time Relations

R.D. Ingthorsson
Lund University and Linköping University, Sweden
rognvaldur.ingthorsson@fil.lu.se

Abstract

This paper is a partial defence of presentism against the argument from cross-time relations. It is argued, first, that the Aristotelian view of causation and persistence does not really depict these phenomena in terms of relations between entities existing at different times, and indeed excludes the possibility of such cross-time relations obtaining. Second, it is argued that to reject the existence of the past—and thereby be unable to ground the truth of claims about the past—does not lead to any absurd consequences.

Keywords: Presentism, cross-time relations, grounding objection, causation, persistence.

1 Introduction

Presentism—the view that only the present exists, the future not yet, the past no longer—is currently a focal point of contemporary philosophy

of time. It is still by no means the favoured view, and yet undeniably an equally hot topic among those who promote it (see, for instance, Crisp 2005, [10]; Bourne 2006, [4]; Markosian 2004, [22]; Ingthorsson 2017, [17]), and those who oppose it (see for instance, Oaklander 2010, [26]; Mozersky 2015, [24]; Torrengo 2017, [34]).

I am not entirely sure why presentism is getting all this attention. It may be related to the growing realisation that presentism is immune to *McTaggart's Paradox* and *the problem of temporary intrinsics* (Craig 1998, [8]; Cameron 2015, [6]; Ingthorsson 2016, [16]) and therefore emerges as the most promising version of the *A* view of time. As a consequence contemporary philosophy of time gravitates towards presentism, and in particular its weak spot, which is undeniably the *problem of cross-time relations* (Bigelow 1996, [3]; Crisp 2005, [10]). I suspect that a contributing factor is also a growing dissatisfaction with the *B* view of time, i.e. its failure to convincingly accommodate for enduring particulars and more generally to account for the dynamic features of experience (Prosser 2016, [28] is possibly the best attempt yet).

The *argument from cross-time relations* goes as follows. If the future/past do not exist, then relations popularly believed to hold between the present and the future/past, do not exist either. Since cross-time relations have come to figure centrally in the formulation of a range of metaphysical notions, like *causation* and *persistence*, and metaphysical cum semantic relations such as *truth* (albeit only the truth of propositions about the past and future), truth being central to our understanding of *knowledge*, then presentism is taken to entail the absurd conclusion that there is no causation, nothing persists, no claims about the past/future are true, and therefore we have no knowledge of the future/past.

The argument from cross-time relations would be a *reductio ad absurdum*, *if* presentists were unable to offer viable alternatives, but they can. In this paper I will first present contemporary versions of presentism and how they propose to deal with the problem of cross-time relations (Section 2). Then I will paraphrase what I have argued elsewhere, notably that a presentist can accept that there can be no true propositions about the past and future, and yet argue that we can have justified belief about the future and past (Section 3). Whether that is enough for having knowledge, is a controversial issue, but it seems to me that this already is a controversy even for that part of temporal reality whose ex-

istence we all agree about—the present—and even if it be assumed that the future and past exist in parity with the present. The controversy, briefly, is that if knowledge is assumed to be fallible, then this is to assume that beliefs about the world can count as knowledge even if they are false. If only those beliefs count as knowledge that are actually true and justified, we have decided to confine the sphere of knowledge to things that are infallible.

Finally, I will argue that the presentist can easily provide an alternative account of causation and persistence, one that does not involve cross-time relations (Section 4). They need not invent one from scratch, because they can make use of the old Aristotelian conception of causation, which did not represent causation as a two-place, cross-temporal, relation. Indeed, most neo-Aristotelian powers-based accounts—which depict causation as the production of change through the interaction of powerful particulars—are already viable presentist accounts of causation (Ingthorsson 2002, [14]).

2 Presentism

In the contemporary literature, presentism is seldom described by more than a single phrase such as 'only present objects exist', or 'existence is confined to the present', before the discussion turns to the problems of upholding that belief. Authors often make up for the brevity of their presentation by appealing to the reader's intuitive understanding of the thesis, claiming that the confinement of existence to the present is a part and parcel of the lived experience of every human being (Bigelow 1996, [3]).

Nevertheless, one can discern a number of alternatives, all of which arise from concerns about the reference and truth of expressions about the future and past. Indeed, all of them deploy what I have elsewhere called the *relocation strategy*; to argue that the entities we naively believe to be in the past or future, are actually in the present (Ingthorsson 2017, [17]). I will focus on the past, since the lack of truth-values for past tensed expressions is a much greater problem than the lack of truth-values for future tensed expressions.

GLOBAL PROPERTY PRESENTISM

One version of presentism, sometimes called *global property presentism* (Kierland 2013, [20]), is usually traced back to Arthur Prior's suggestion that claims about the past, even those that have the grammatical structure of singular statements, do not refer to any particular existing state of affairs, but to some general feature of the world as it is now:

> the fact that Queen Anne has been dead for some years is not, in the strict sense of 'about', a fact about Queen Anne; it is not a fact about anyone or anything—it is a *general* fact. Or if it is about anything, what it is about is not Queen Anne—it is about the earth, maybe, which has rolled around the sun so many times since there was a person who was called 'Anne', reigned over England, etc. (Prior 1962, p. 13, [27])

Please note that by 'fact' Prior means something like true proposition, while whatever facts are 'about' refers to the existent state of affairs that makes that proposition true. Anyway, for Prior, facts about the past constitute the set of beliefs that are putatively true and thus in need of truthmakers. The problem is that the past does not exist, according to Prior's presentism, and hence he tries to find plausible candidates in the present to act as referents and truthmakers for claims about the past. Prior admits he isn't too sure what exactly these the truthmaking features are, and suggests that it could even be some feature of the earth as a whole, which is why this position is called global property presentism.

Prior's point is not grammatical but ontological. He is not arguing that the true grammar of past tensed expressions implies that we are really talking about general features of the earth. I am sure he recognises that the grammar of 'Queen Anne is dead' is most naturally read as the attribution of something to some particular entity. But his appraisal of the ontology of time, i.e. that the past does not exist, convinces him that the grammar of past tensed expressions *must be misleading*, and consequently suggests another way of understanding them.

Prior's position can be criticised for being too vague. Surely, the claim 'Queen Anne's death was peaceful' and 'dinosaurs roamed the Earth' are made true by different things, but his suggestion doesn't give much guidance as to which features of the present this would be. Very

plausibly, Prior means to say that Queen Anne's death must have left a mark on reality, to contribute somehow to what it is like today. However, he doesn't give any details about how such marks are left, and where to find them. I'll venture to offer more detail. We find out about Queen Anne's death by checking what is documented in historical records. So, the *historical record*, and its endurance in the present from the time of her death until now, is a much better suggestion than the earth as a whole about what it is about the present world that justifies our belief that there indeed was such a things as Queen Anne's death. Fossilised remains of dinosaurs represent the particular marks left on reality by the dinosaurs of the past, preserved in what could be called the *natural record* of the past. We then have two distinct features of the present world to support our belief in the death of Queen Anne and in the prehistoric existence of dinosaurs. But do such entities really make our beliefs true, as opposed to simply constituting our grounds for belief? Well return to that question later.

ABSTRACT ENTITY PRESENTISM

Others have suggested, as truthmakers for past tensed expressions, various abstract entities that are not to be understood as properties of the concretely existing present. For instance, that expressions like 'Socrates was wise' refers to an *individual essence* of Socrates, a kind of abstract entity that could exist even if Socrates' has physically ceased to exist (Craig 2000, p. 199, [9]). I find this suggestion too arbitrary and epistemically uninformative. It's very easy to postulate that everything that comes to be in the present leaves an incorporeal trace of itself that somehow endures in the present thereafter—and to appeal to such incorporeal entities to support one's belief that P is true—but it is difficult to identify those traces and extract any information about the past from them. Furthermore, this strategy converts presentism from one of the sparsest ontological doctrines of all, into something much less sparse; in addition to the present being constituted by the current state of the world, it also consists in the incorporeal individual essences left by every becoming in the history of the universe. I don't have a knock down argument against such a view. The best I can do is to challenge the need to postulate it.

Another option is *ersatz presentism*. Instead of postulating individual abstract entities for each concrete particular and/or event that ever existed, ersatz presentism postulates that at each present moment, there exists, for every future and past time, a set of propositions representing the state of the world at those times (Bourne 2006, [4]; Crisp 2007, [11]). Now, sometimes it appears that such ersatz times are only postulated for representational purposes, to allow us to talk about the future and past (Wüthrich 2012, p. 445, [35]), but of course this would not go a long way towards answering any objections to presentism. On the other hand the same writers claim that ersatz presentism represents a promising solution to the problem of reference and truth (Wüthrich 2012, *abstract*, [35]).

Like Mozersky (2015, p. 44, [24]), I worry that if ersatz times are assumed to have a truthmaking function this threatens to reverse the ontological priority of the present. The core idea of presentism, as I understand it, is that reality is grounded in the concretely existing present. Ersatz presentism instead construes times as abstract propositions, which then somehow make true anything expressed about concrete events at various presents. Now, obviously it is possible to argue that the ersatz times that represent the past are marks left by the past, in the same way individual essences are, but this won't work for the future, because it has yet to make a mark on the present.

In addition, ersatz presentism doesn't explain either how these ersatz times figure in our epistemic practices and so cannot be used to support our beliefs about the past. It is still the case that we learn about dinosaurs from fossils in the ground and not by accessing a realm of abstract propositions. Furthermore, none of the options that postulate abstract entities address at all the issues we have about causation and persistence.

The main worry about these general, incorporeal, and abstract entities that presentists are postulating is that they are entities that do not really come with a story of genesis; how do they come into being, if at all? Nor do they come with a story about how they figure in our epistemic practices. They are difficult to understand as marks on present reality that could somehow *inform* us about the past. They appear instead as arbitrary postulations about the world having whatever feature required to make some or other expression about the past true. In other words, the preferred procedure seems to be that we first decide—on sheer in-

tuition it seems—which propositions about the past are true, to then infer that the world must at present bear some truthmaking features that make those propositions true. This just isn't in conformity with what we actually do in our epistemic practices. We justify our beliefs about the past by appealing to the historical and natural record. These records exist concretely in the present in the form of various man made artefacts (books, audio recordings, photos, films, archaeological artefacts, *etc*.) and natural remains such as fossils, layers of soil and rock, oil, *etc*. Nobody says that Socrates was wise because this is evident from his individual incorporeal essence, but because of what we learn from the historical record.

LUCRETIAN OR NOMIC PRESENTISM

A more concrete suggestion about what counts as a mark of the past in the present, is the concrete and determinate state of the world at any given time. The basic idea is that when things are alive and kicking, they leave a concrete mark on enduring reality; we have scars left as reminders of former times, and we find fossils in the ground. John Bigelow tells us that this is a core idea in Stoic presentism, and he cites Sextus Empiricus as saying: "if this man has a scar, this man has had a wound" (1996, p. 41, [3]). If we perceive that a man has a scar *now* we can infer that he has had a wound in the past. It isn't clear to me whether Empiricus intends to say that we actually ascribe the property of 'having had a wound' to the scarred man, but at least he is saying that we can infer from the state of things now, how something was in the past. This satisfies my qualms about how we find out about the past. The question is if our grounds for believing p, must be the same as what actually makes p true.

The name 'nomic presentism' (Kierland 2013, [20]), as far as I can tell, really denotes the same basic idea, but is more concerned with the modern idea that that on the basis of our knowledge of the qualitative state of the world at any given present, and our knowledge of the laws of nature, we are able to infer what the world was like in the past (and predict how it will be in future). To do this we don't need to postulate the existence of incorporeal, general, or abstract entities existing somehow parallel to the existing qualitative state of concrete reality. The laws of

nature are in turn perfectly befitting a presentist ontology, because the world instantiates them at any given time.

Now, I like nomic presentism as a view of the world and of how we find out about the world, but not as a way to make our claims about the past true. First, as a thruthmaking theory it requires the world to be causally determined, in order that each state of the universe passes on—in an unbroken chain—information not just about itself and each immediately preceding state, but about every preceding stage. If the world was like that, it would certainly allow us to extract information about the past, but also makes the future as fixed and determined as the past. I find this to be an unwanted consequence.

The second problem is that even if we grant causal determinacy and ignore the problems of a fixed future, I still don't see how the present + laws can ground the truth of our beliefs about the past; not on any extant theory of truth. The present + the laws of nature, doesn't look in any way like the past, so it can hardly make our beliefs about the past true by corresponding to them. Appeal must be made to a completely different notion of truth.

There are theories about truth that do not require the existence of whatever it is a proposition is about, but I can't see that they will do the nomic presentist any good. The coherence and pragmatic theories of truth do not make truth reliant on what the world is like at all. The deflationary theory arguably rejects truth altogether, and the identity theory construes truth as an identity of a belief to a true proposition, not to the world (for a more detailed discussion of various truth-theories, see Ingthorsson, *forthcoming* [18]). The identity theory may work for some forms of erzats presentism, but not nomic presentism. The suggestion will only work if we understand 'making true' in some altogether new and primitive way. I have yet to see presentists take a stand on this issue and won't speculate further on this issue here. Instead I will consider the alternative most presentists do not consider at all; why not simply deny that expressions about the past have truth values? I for my own part am perfectly satisfied as long as it is possible for a presentist to *justify* her beliefs about the past, say, that dinosaurs roamed the earth.

3 True vs. Justified Beliefs about the Past

Presentists deny the existence of the past, but they acknowledge the existence of everything that any scientific discipline has ever *de facto* appealed to in their justification of theories/hypotheses about the past—man made documentation, fossils, evolution theory, knowledge of the laws of nature—because they all obtain now. Consequently, justification is obviously no problem at all. But, the beliefs we justify about the past cannot correspond to anything. Does this mean that we cannot say that we have knowledge about the past? Well, it does at least imply that our knowledge about the past is at best hypothetical and fallible, which is what is already acknowledged about our knowledge about the present. If knowledge is justified belief that is also true, knowledge is by definition *infallible*. A belief that is true, will not just never happen to be falsified; it cannot be falsified (it can only be falsely falsified by some experimental mistake).

More than anything, this problem revolves around the philosophical question of how exactly to understand knowledge. It does *not* really concern the epistemology of the past or our current epistemic practices. No A- or B-theorist argues that we find out about the past in any other way than by inferring it from the historical and natural record as it exists at any given time. That simply is the way science works. Accordingly, this discussion only concerns our general attitude towards the past—do we believe it exists or not—and with our understanding of knowledge generally speaking. I will focus on the latter question, which I understand to be a question of whether we can allow knowledge to come in degrees; is some knowledge both justified and true, and is some knowledge only justified?

The first thing to note that a conception of knowledge about the past as only justified but not true, is not equal to a conception of knowledge about the present as only justified but not true. If we believe something about the present on the basis of the available evidence, but it fails to correspond to reality because the available evidence was incomplete, then our belief is false because it represents reality as it really isn't. However, if we believe something about the past on the basis of the available evidence but it fails to correspond to reality because the past has ceased to exist, this does not mean that the belief represents the past as it really wasn't. Sure, to make this argument really stick, one would

have to develop a theory of falsemaking, which I will not do here. However, I think the point is intuitively clear enough for my present purposes, notably that lack of truthmakers for past tensed expressions does not make our ideas about the past into misrepresentations of the past, but the lack of truthmakers for present tensed expressions will inevitably mean that these expressions misrepresent reality.

Second, the idea that knowledge has to be true *and* justified is already too strict to comply with received views about what counts as knowledge. We generally call everything knowledge that strikes us as justified on the basis of the available evidence, never mind whether it actually is true. We even call some things knowledge that we know is false. Take classical mechanics as an example. We know that classical mechanics is at best a useful approximation to reality, but it continues to be a staple in physics education and continues to be called knowledge. It continues to be a part of the curriculum because it is so useful and much easier to apply in the situations where it is known to give the same results as quantum mechanics and theory of relativity respectively. At the very least, our beliefs about the past that are justified by the historical record would continue to be called knowledge even if we agree it cannot correspond to a past that no longer exists. Otherwise put, the concept of knowledge that is already in use, is one that allows of degrees.

On a related note, some may worry that the lack of truth-values for expressions about the past implies that the past is indeterminate. Statements about the past are popularly believed to have determinate truth-values because once things happen in a certain way in the present there is no way to undo or change it. However, I can't see that truth has much to do with determination. For mind-independent reality to be determinate, it is not required that there be propositions about it with determinate truth-values, nor is it required that the past be determinate for it to be true that whatever happens in the present is determinate and can never be undone. It is enough to know that what happens in the present is always determinate for us to know that what happened in the past also *was* determinate, because when it happened it was present and thus determinate. This conclusion holds whether or not we know anything about the past. The intuition that once things have happened, they can never be undone, is satisfied perfectly well by the consideration that once things have happened and ceased to exist, they

cannot be undone; you cannot go back to a non-existent past to undo it. And anyway, it is supposed to be the case that propositions have determinate truth-values because reality is determinate, not the other way around.

So, would it be so outrageous to suggest that the idea of knowledge being true justified belief may perhaps be useful as a regulative idea for what we ideally strive towards, but fails to demarcate between what is today counted as knowledge and what doesn't. All things considered, it would seem a little thing to allow our beliefs about the future and past to be called knowledge, if the available evidence justifies them, even if they cannot in principle correspond to anything. I take it to be an open question still, in epistemology, whether knowledge is to be demarcated in terms of true justified belief or not. Consequently, philosophy of time should not decide in favour of one particular theory of time on the basis of a premature stance on what is the correct view of knowledge. That would be a case of letting one epistemic position override any metaphysical concerns, to settle a metaphysical issue.

What I have so far argued with regards to knowledge and truth does not of course prove anything. It just serves to show that the consequences of denying that past and future tensed propositions can be true, are not so serious as to make that option unthinkable. It does not have any implications for our current epistemic practices, nor diminish our prospects of justifying our beliefs. Let me now turn from truth and knowledge, to persistence and causation.

4 Persistence and Causation are Not Cross-Time Relations

Can persistence and causation be explicated in presentistic terms, i.e. without invoking cross-time relations? The answer is yes, and this should not be news to anyone. The Aristotelian account of change, persistence, and causation already is presentistic. Let us briefly consider the Aristotelian account of change, in contrast to the contemporary characterisation of change as "simply difference or nonidentity in the features of things" (Mortensen 2016, sect. 1, [23]). The latter is sometimes cashed out more formally in terms of a conjunction of (or difference between) two states, i.e. 'a-is-G-at-t' and 'a-is-*not*-G-at-$t*$', which can

very easily be interpreted as a cross-temporal characterisation of change, notably as a relation between the temporal parts of *a* which are *G* and *not-G* respectively. This is what Johanna Seibt calls the 'state analysis of change', which she argues is one of many unquestioned presuppositions in what she calls the 'paradigm of substance ontology' (1996, [31]). But the substance ontology she has in mind is not the Aristotelian ontology; it is the ontology favoured in 20th Century analytic philosophy, by philosophers that notoriously resist primitive notions.

Far from being a relation across times between two states of a thing, the Aristotelian account of change in terms of alteration, i.e. a material substance ceasing to manifest one quality and beginning to manifest another contrary quality, *excludes* that change can involve any such relation. Since the two states of a thing, before and after the change, are contrary states of one and the same entity that *endures* through the change, the existence of one state excludes the existence of the other. Change simply cannot be a relation between existent states, located (existing) at different times, at least if we accept the idea that relations can only hold between existent entities.

As I have argued elsewhere (2001, [13] and 2016, ch. 7, [16]) the problem of temporary intrinsics—which is meant to show that things cannot really endure—arises only when it is first assumed that all times exist in parity; it is only on the assumption that '*a*-is-*G*-at-*t*' and '*a*-is-*not*-*G*-at-$t*$' are equally existent and real entities located at different times that we get the conclusion that *a* is equally *G* and *not-G*. Indeed, David Lewis admits that presentism, which denies temporal parity, avoids the problem (1986, p. 222ff, [21]). Accordingly, the only thing the problem of temporary intrinsic establishes is that the attempt to combine endurance and eternalism leads to contradiction; ergo, things cannot endure in tenseless time (for a more detailed argument, see Ingthorsson 2009, [15]).

Otherwise put, the Aristotelian account does not portray a succession of states as *constituents* of change—it is not what change consists in—but a *consequence* of change, i.e. of the alteration of something from one state to a contrary state while remaining numerically the same. The state analysis can really only be understood either as a description of the *appearance* of change (we first observe *a* to be *G*, and later we observe it to be *not-G*) or it is a statement of what change must be like if one assumes eternalism to be true.

Indeed, persistence, and causation cannot possibly be cross-temporal relations either on the Aristotelian account, for the very same reason change cannot be a cross-temporal relation (admittedly, this is not as obvious in the case of causation). Persistence cannot be a cross-temporal relation if, as the Aristotelian assumes, things persist by enduring; i.e. if they pass from one time to another and in that process cease to exist at the time it passes from. On this view, an object existing wholly at t_1 cannot stand in a relation to itself at other times because it doesn't exist at any other time. Indeed, already Aristotle addressed the problem of temporary intrinsics, which he attributed to the Sophists, notably the argument that Chrysippus in the market place is not identical to Chrysippus in the gymnasium, because some things hold true of the former that does not hold for the latter (*Physics:* Bk. 4, Part 11, [2]). Aristotle's solution is that Chrysippus remains numerically the same while loosing and acquiring properties as he saunters from the market place to the gymnasium. When in the market place, there is no Chrysippus in the gymnasium, and vice versa.

Another interesting example of when prior commitments affect the appreciation of philosophical views—and which helps to understand why contemporary philosophers are not comfortable with presentism—is that it is difficult to deal with presentism in the language of first order predicate logic. At least if first order predicate logic is meant to function like Quine prescribed, notably to specify our existential commitments (for arguments to this effect, see Øhrstrøm and Schärfe 2004, [36]; Seibt 2016, [33]). Sure one can introduce temporal operators, but on Quine's understanding, such operators must operate on something existing, wherefore the use of past tense operators to talk about Chrysippus in the market place is still to quantify over existents (there exists an x such that Px). Indeed, as Øhrstrom and Schärfe argue, it was concerns about Quine's idea about ontology that drove Arthur Prior to develop a temporal logic of a different kind.

Now, I do not want to get entangled in the details of the endurance vs. perdurance debate. I have nothing to add to what I have elsewhere argued (2001, [13]; 2009, [14] and 2016, ch. 7, [15]). The important point for this paper is simply to point out that the problem of cross-time relation is not a *reductio ad absurdum* of presentism because alternative explanations of persistence and causation are available. The alternatives may well be problematic in many respects, but those who appeal to the

problem of cross-time relations do not often take any such problems into consideration; they typically assume that no options exist.

Let us now consider causation, and again turn to Aristotle, at least initially. As with the case for change, there is a tension between the Aristotelian account of causation and what I take to be the received view in philosophy today, notably, that causation is at rock bottom a relation between temporally distinct events. The latter has become so well entrenched in the philosophical tradition that there is little or no awareness of alternatives. Consider that Jonathan Schaffer's entry on 'The Metaphysics of Causation' in *The Stanford Encyclopedia of Philosophy* (2016, [30]) is written entirely on the assumption that all extant views portray causation as a relation, and that the controversies about causation only revolve around the nature of the relata or of the relation. Schaffer does discusses various problems with what he calls *causal processes*, but deals with them as if they are composed of sequences of stages and/or events. Schaffer seems unaware of the fact that philosophers like Salmon (1984, p. 139ff [29]) and Seibt (2000, [32]), explicitly reject an event ontology of processes, and likewise that Ingthorsson (2002, [14]), Chakravartty (2005, [7]) and Mumford and Anjum (2011, [25]) explicitly argue that causation should be thought of as a process *instead of* as a relation. Clearly there are non-relational conceptions of causation also in contemporary philosophy, but let us here focus on the traditional Aristotelian view.

Note however, that the account I offer is not Aristotle's original view. It is a paraphrase of a view widely attributed to the Aristotelian school of thought. It comes very close to the account stated by Hobbes in the very early beginnings of empiricism (1656, ch. IX–X, [12]). Bunge (1959, ch. 2, [5]) and Johansson (1989, p. ch. 12, [19]) offer similar paraphrases of the causal realist tradition they trace back to Aristotle. The main difference from the original view, is that it does not include final causes. Indeed, already Hobbes argued against final causes (1656, p. ch. X, sect. 7, [12]), although his account is roughly Aristotelian in many other respects

According to the roughly Aristotelian view I have in mind, then, new states of affairs are produced when an already existing material body, or complex of bodies, changes due to an external influence without which the change would never have come about and the new state of affairs never exist. The kernel of this view comes out clearly in the slogan

'whatever comes to be is necessarily born by the action of a cause'; very probably a paraphrase of Aristotle's claim that "everything that comes to be comes to be by the agency of something and from something and comes to be something" (*Metaphysics*, bk. 7, part 7, [1]). Typically, the external influence, or cause, is depicted in terms of an 'extrinsic motive Agent' (or, simply *Agent*, i.e. an object possessing an active causal power), which exerts that power upon another object. The latter objects is typically called *Patient* since its role in the interaction is to passively receive the influence exerted by an Agent and change in some specific way in accordance to its passive power, i.e. an ability to change in some specific way in response to the influence of the active power. Accordingly, a *cause* is the exertion of influence by an Agent upon a Patient and an effect is the resulting change in the Patient.

When I say that a cause is the action of an Agent upon a Patient, then one should not understand a cause as merely the action of the Agent, but as the *interaction* between Agent and Patient. It is this interaction that I identify with a *process of production*, and there are two salient features of this process that need emphasis, because they stand in stark contrast to the relational view of causation. The first point is that *causal influence* is something that is exerted by an Agent on a Patient. In other words—and this is critical for understanding the main point of this paper—actions occur between *persistent objects*, not between *events* or *states*. It is not the cause that acts on or influences the effect, indeed, that is impossible. If the effect only comes into existence by being produced by the cause, i.e. by the action of the 'efficient cause', the effect cannot be subject to the very same action that is supposed to produce it. To assume the effect is subject to the action that produces it, is to assume the effect already existed when it is acted upon, and thus could not have been produced by that very same action. Indeed, we see in the following passage from Hobbes a clear statement both of the idea that actions occur between persistent objects, and that the effect only comes into being as a consequence of actions between persistent objects:

> A body is said to work upon or *act*, that is to say, do something to another body, when it either generates or destroys some accident in it: and the body in which an accident is generated or destroyed is said to suffer, that is, to have something done to it by another body; as when one body by put-

> ting forwards another body generates motion in it, it is called an AGENT; and the body in which motion is so generated, is called the PATIENT; so fire that warms the hand is the *Agent*, and the hand, which is warmed, is the *Patient*. That accident, which is generated in the Patient, is called the EFFECT (Hobbes 1656, part II, ch. IX, sect 1, [12])

The second point is that the Aristotelian view depicts effects as the product not of the action of the Agent alone, but of a *total cause* of a certain kind; it is a product of the way two or more material bodies act on each other in virtue of their powers to produce a change in those very bodies. Again we can appeal to Hobbes as witness:

> [...] an entire cause, is the aggregate of all the accidents both of the agents how many soever they be, and of the patient, put together; which when they are all supposed to be present, it cannot be understood but that the effect is produced at the same instant; and if any one of them be wanting, it cannot be understood but that the effect is not produced.
> (Hobbes 1656, ch. X, sect. 3, [12])

Joining now together the Aristotelian account of change, persistence, and causation, there is not much conceptual space left to think of causation as a two-place relation between two existents existing at different times. Whatever material entities there exist at one time will pass in their entirety to the next by virtue of enduring, whether it be unchanged or changed by any occurring causal influence. Furthermore, an event or state existing at one time does not cause the next by somehow standing in an unanalysed and/or primitive relation of 'production' to the later event/state. Causal production is analysed in terms of a change in a complex of objects provoked by an influence exerted between those objects. Since the influence is exerted between objects, and the temporal relation between those objects is permanently synchronous, there never is a diachronic relation of influence between anything. In particular, the relation of production cannot be such a diachronic relation because whatever produces anything ceases to exist in that process.

The Aristotelian view I have described seems to me to be in accordance to all major schools of thought before the rise of empiricism (in

the particular respects considered here). In Atomism, Stoicism, Scholasticism as well as in the corpuscular view endorsed by the natural philosophers of the early enlightenment, the common assumption is that influence is exerted between two material objects whose relation is synchronous. Indeed, as far as I can tell, this is still a standard understanding in particle physics. All the fundamental forces of nature are exerted between persistent entities (some of them do persist very briefly, that is true, but persist nevertheless). In the *Large Hadron Collider* they are not accelerating events to make them smash into other events; they accelerate particles to make them smash together to break each other up. In chemistry the assumption is that various substances react with each other. Oxygen reacts with some fuel to combust; oxygen does not react with combustion. Furthermore, the common sense conception is that bodies act on each other: the leaden ball dropped upon a pillow acts on the pillow to make a hollow (the falling doesn't act upon the forming of a hollow); the horse pulls the cart (not: the motion of the horse that pulls the motion of the cart); the brick hits the window (not: the motion of the brick hits the breaking of the window).

5 Conclusion

The very humble conclusion that this paper leads up to, is simply that presentists would not be absurdly out of touch with reality were they to suggest that future and past tensed propositions just are not true, or to deny that persistence and causation are at rock bottom cross-time relations. They should argue that while future and past tensed expressions just can't be technically true, we still have all the reasons we ever have had to believe what the past used to be like and what the future will be. No loss epistemically to deny the existence of the future and past. And they should point out that the idea that persistence and causality are cross-temporal relations is a peculiarly Humean/empiricist conception that doesn't resonate with the Aristotelian view of things.

Acknowledgements

This paper is based on research in the project "Scientific Essentialism: Modernising the Aristotelian View", funded 2015–2018 by *Riksbankens*

Jubileumsfond: Swedish Foundation for Humanities and Social Sciences, Grant-ID: P14-0822:1.

Bibliography

[1] Aristotle (*Metaphysics*). *Aristotle's Metaphysics*. Ross, W.D. (ed.). Oxford: Clarendon Press, 1924.

[2] Aristotle (*Physics*). Physics. In: Wickstead, P.H. and Cornford, F.M. (transl.): *Aristotle*. Vol. IV. Cambridge Massachusetts: Harvard University Press, 1957.

[3] Bigelow, J. (1996). Presentism and Properties. *Philosophical Perspectives* **10**: 35–52.

[4] Bourne, C. (2006). *A Future for Presentism*. Oxford: Oxford University Press.

[5] Bunge, M. (1959). *Causality*. Cambridge Massachusetts: Harvard University Press.

[6] Cameron, R. (2015). *The Moving Spotlight: An Essay on Time and Ontology*. Oxford: Oxford University Press.

[7] Chakravartty, A. (2005). Causal Realism: Events and Processes. *Erkenntnis* **63**(1): 7–31.

[8] Craig, W.L. (1998). McTaggart's Paradox and the Problem of Temporary Intrinsics. *Analysis* **58**: 122–7

[9] Craig, W.L. (2000). *The Tensed Theory of Time*. Dordrecht: Kluwer Academic Publishers.

[10] Crisp, T.M. (2005). Presentism and "Cross-Time" Relations. *American Philosophical Quarterly* **42**(1): 5–17.

[11] Crisp, T.M. (2007). Presentism and the Grounding Objection. *Noûs* **41**: 90–109.

[12] Hobbes, T. (1656). *Elements of Philosophy Concerning Body.* In: W. Molesworth (ed.): *The English Works of Thomas Hobbes of Malmesbury*. London: Bohn, 1839.

[13] Ingthorsson, R.D. (2001). Temporal Parity and the Problem of Change. *SATS–Northern European Journal of Philosophy* **2**(2): 60–79.

[14] Ingthorsson, R.D. (2002). Causal Production as Interaction. *Metaphysica* **3**(1): 87–119.

[15] Ingthorsson, R.D. (2009). Can Things Endure in Tenseless Time? *SATS–Northern Europe Journal of Philosophy* **10**(1): 79–99.

[16] Ingthorsson, R.D. (2016). *McTaggart's Paradox*. New York: Routledge.
[17] Ingthorsson, R.D. (2017). Challenging the Grounding Objection to Presentism. *Manuscrito* **40**(1): 87–107.
[18] Ingthorsson, R.D. (forthcoming). There's No Truth-Theory Like the Correspondence Theory. *Discusiones Filosóficas*.
[19] Johansson, I. (1989). *Ontological Investigations: an Inquiry into the Categories of Nature, Man*, and Society. New York: Routledge.
[20] Kierland, B. (2013). Grounding Past Truths: Overcoming the Challenge. In: Ciuni, R., Miller, K. and Torrengo, G. (eds.): *New Papers on the Present—Focus on Presentism*. Philosophia Verlag: pp. 173–209.
[21] Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Basil Blackwell.
[22] Markosian, N. (2004). A Defense of Presentism. *Oxford Studies in Metaphysics* **1**(3): 47–82.
[23] Mortensen, C. (2016). Change and Inconsistency. In: Zalta, E.N. (ed.): *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archves/win2016/entries/change/. Accessed Nov 12, 2018.
[24] Mozersky, J.M. (2015). *Time, Language, and Ontology*. Oxford: Oxford University Press.
[25] Mumford, S. and Anjum, R.L. (2011). *Getting Causes from Powers*. Oxford: Oxford University Press.
[26] Oaklander, L.N. (2010). McTaggart"s Paradox and Crisp"s Presentism. *Philosophia* **38**(2): 229–41.
[27] Prior, A.N. (1962). *Changes in Events and Changes in Things*. Lawrence, Kan.: University of Kansas Press.
[28] Prosser, S. (2016). *Experiencing Time*. Oxford: Oxford University Press.
[29] Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. New Jersey: Princeton University Press.
[30] Schaffer, J. (2016). The Metaphysics of Causation. In: Zalta, E.N. (ed.): *The Stanford Encyclopedia of Philosophy.* https://plato.stanford.edu/archives/fall2016/entries/causation-metaphysics/. Accessed Nov 12, 2018.
[31] Seibt, J. (1996). Existence in Time: From Substance to Process. In: Faye, J., Scheffler, U. and Urs, M. (eds.): *Perspectives on Time*.

Boston Studies in Philosophy of Science. Dordrecht: Kluwer: pp. 143–82.

[32] Seibt, J. (2000). The Dynamic Constitution of Things. *Poznan Studies in the Philosophy of the Sciences and the Humanities* **76**: 241–78.

[33] Seibt, J. (2016). Process Philosophy. In: Zalta, E.N. (ed.): *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/win2017/entries/process-philosophy/. Accessed Nov 12, 2018

[34] Torrengo, G. (2017). The Myth of Presentism's Intuitive Appeal, *Phenomenology and Mind* **12**: 50–56.

[35] Wüthrich, C. (2012). Demarkating Presentism. In: de Regt, H. Okasha, S. and Hartmann, S. (eds.): *EPSA Philosophy of Science: Amsterdam 2009*. Dordrecth: Springer: pp. 441–50.

[36] Øhrstrøm, P. and Schärfe, H. (2004). A Priorean Approach to Time Ontologies. In: Wolff, K.E. et al. (eds.): *International Conference on Conceptual Structures 2004: Conceptual Structures at Work*, Dordrecht: Springer: pp. 388–401.

Ingthorsson, McTaggart's Paradox and the R-theory of Time

L. Nathan Oaklander
University of Michigan-Flint, USA
lno@umflint.edu

Abstract

In his provocative book, *McTaggart's Paradox* (2016, [9]), R.D. Ingthorsson argues that McTaggart's argument for the unreality of time rests on the principle of temporal parity according to which all times or events in time exist equally or co-exist in a sense that is compatible with their being successive. Moreover, since temporal parity is also an essential tenet of the B-theory, McTaggart's argument against the reality of time can also be used to undermine the B-theory. Ingthorsson argues further that only by adopting an ontologically frugal presentist metaphysics can one avoid McTaggart's paradox and account for identity through time and change. The aim of this paper is to clarify Russell's authentic view of time in contrast to the B-theory which is McTaggart's misrepresentation of Russell and argue that temporal parity it is not a fundamental tenet of the Russellian (R-) theory. For that reason, the R-theory is immune to objections that are based on temporal parity. I shall then offer my own interpretation of McTaggart's paradox that renders Ingthorsson's version of presentism subject to it.

Keywords: McTaggart's Paradox, A-series, B-series, temporal parity, Russellian (fact) ontology, substance ontology, presentism.

1 Introduction

In his valuable book on *McTaggart's Paradox*, R.D. Ingthorsson (2016, [9]) makes some provocative claims regarding the scholarship on McTaggart's argument for the unreality of time. Most notable is his claim that the key assumption that McTaggart employs in his argument, namely, the principle of temporal parity—the view that all times (whether A-times or B-times) exist equally or co-exist—is a basic tenet of the B-theory and therefore, McTaggart's paradox cannot be used to support the B-theory. In other words, McTaggart's argument for the unreality of the A-series rests on a premise that also undermines the B-series, and the B-theory of time. Ingthorsson expresses this point as follows: "But, in the end, the central question still is, do all times exist in parity or not? If they do [as the B view maintains], then how should we understand such a reality as temporal? That is, wherein lies the temporality of earlier than and later than?" (Ingthorsson 2016, p. 141, [9]).

A further provocative claim Ingthorsson makes is that virtually all commentators on McTaggart's philosophy of time have failed to recognize the connection between his *a priori* metaphysics, that is, his view of Absolute Reality, and his view of Present Experience. Ingthorsson maintains McTaggart's argument for time's unreality is intended to demonstrate that time as it is given to us in experience, as stated in the chapter "Time," in (1927, [14]), is incompatible with his view of Absolute Reality found in (1921, [13]). For that reason, Ingthorsson is highly critical of Broad (1933, p. 9, [3]) who claims that McTaggart's paradox is a stand-alone argument that is not dependent on the results in the first part of his inquiry.

Another controversial thesis of Ingthorsson's book is that McTaggart is not concerned with the language of time, but the ontology of time. McTaggart's argument is not based on the proper semantic analysis of sentences which record the facts of temporal becoming, such as "It *is* raining," and "It's raining *has been future* and *will be* past," but with their ontological analysis. Ingthorsson argues that given the principles that underlie McTaggart's *substance ontology*, the ontological ground of time as we experience it and express it in ordinary language cannot exist, and since McTaggart believes that only what exists is real, he concludes that time is unreal. Thus, it is a mistake to claim as Broad (1938, p. 317 [4]), Lowe (1987, [11]) and others have done that McTag-

gart's argument is based on linguistic confusions.

At the outset, I would like to say that I applaud Ingthorsson's emphasis on the ontology that underlies McTaggart's argument, and given the ontology he attributes to McTaggart, I would not contest that he demonstrates that time is unreal. So, I wholeheartedly agree with his emphasis on the ontology and not the language of time. However, I shall argue that a textual case can also be made for questioning the claim that McTaggart's *a priori* metaphysics developed in (1921, [13]) is what he employs to demonstrate the unreality of time in (1927, [14]). There is an important difference between the ontology Ingthorsson attributes to McTaggart, and the ontology McTaggart employs to generate his paradox. Indeed, the ontology presupposed by McTaggart's argument (the doctrine of external relations) is inconsistent with his prior commitments. Therefore, not only is Ingthorsson mistaken about McTaggart basing his argument on his prior ontology, I will show that McTaggart is inconsistent in that he bases the argument on an incompatible (Russellian) ontology.

Furthermore, I would also question whether McTaggart's argument for the unreality of time shares a common assumption with *Russell's* view that McTaggart claims to be critiquing in the chapter on "Time," that the B-series alone constitutes time. Ingthorsson claims that just as McTaggart's ontology implies that all temporal positions whether past, present or future co-exist equally even if they are not present, on the B-theory all times co-exist equally even if they are not simultaneous, but successive. Indeed, it is the co-existence of past, present and future that is *common* to McTaggart and the B-theorists. As he puts it,

> To illustrate the fundamental difference between *A* theorists, on the one side, and McTaggart and the *B* theorists on the other [we can say:] The former cannot understand why we should think of the future and past as an existing reality—at least not one separate from what exists in the present—and thus fail to see why we should interpret claims about the future and past as claims about something non-present and yet existing, whereas the latter cannot understand how we can fail to think of the future and past as co-existing with the present, even though they are not co-located in time.
>
> (Ingthorsson 2016, p. 89, [9])

Ingthorsson's thesis is that the only adequate A-theory is presentism, since it rejects temporal parity, whereas all other theories follow McTaggart and B-theorists in accepting temporal parity, and thereby fail. Thus, Ingthorsson maintains that the fundamental debate is not between A- and B-theorists, but between those who adopt the temporal parity thesis (McTaggart and B-theorists and others),[1] and those who deny it (presentists). I shall argue, however, that there is a third alternative that involves a Russellian (or R-theoretic) ontology of time that differs from presentism, McTaggart and the B-theory, as Ingthorsson understands them. The debate between R-theorists and Ingthorsson's version of presentism is between those who view the spatio-temporal universe as the sole purview of reality and those who view the world as greater than the universe, including non-spatial and non-temporal objects. The fundamental debate, in other words is between naturalists and ontologists to use terms that Reinhardt Grossmann (1992, [8]) did to distinguish a fundamental philosophical divide.

To explain what I consider to be wrong with how Ingthorsson characterizes the fundamental agreement between McTaggart and the B-theorists it will be necessary to distinguish McTaggart's understanding of B-relations and the B-series from a Russellian understanding of R-relations and the R-series.[2] Given the incommensurability between McTaggart and Russell's analyses of the "B-series" it is a mistake to believe that the temporal parity thesis in McTaggart's argument against the A- and B-series can also be employed against the R-series to demonstrate that "the conception of temporal reality as a series of events [...] related as earlier/later than each other [...] is a conception of a changeless reality and consequently not a conception of a temporal reality" (Ingthorsson 2016, p. 92, [9]). I shall proceed by briefly summarizing in section II, McTaggart's ontology as Ingthorsson understands it, and how it differs from Russell's ontology. Then, in section III, I will explain key elements of the Russellian or R-theory of time and the principle of temporal parity, demonstrating how they differ. In the fourth section, I discuss McTaggart's paradox and show how it not only undermines the A-series, but that it is also applicable to presentism as Ingthorsson conceives of

[1] Ingthorsson tends to include A-B hybrids, moving spotlight theorists, and even growing blockers among those who accept temporal parity.

[2] I contrast the B-theory and the R-theory below (pp. 9–11). See also, Oaklander, 2016, 2015, 2014a, 2014b and 2012 [17–21].

it. Since his paradox can be interpreted to show that succession, construed R-theoretically, is the ground of the passage of time, and since presentism rejects cross-temporal relations, McTaggart's argument undermines Ingthorsson's presentism as well.

2 McTaggart's Substance Ontology and Russell's Fact Ontology

According to Ingthorsson, McTaggart's metaphysics "is the ultimate substance ontology and alluring in its simplicity. Reality is grounded in the existence of substances that bear properties and stand in relations" (Ingthorsson 2016, p. 15, [9]). Regarding substance McTaggart claims:

> There is a substance which contains all existent content, and of which every other substance is a part. This substance is called the Universe. A universe contains all existent content. Or we can define the Universe as a substance of which all other substances are parts.
>
> (McTaggart 1921, sect. 135, p. 148, [13])

As Ingthorsson interprets him,

> McTaggart presupposes that time, if real, is a compound whole whose every part is equally existent and real. [...] Since existence and reality coincide, and *every existing thing is in time*, then by time he includes the whole of reality. There is no distinction to be drawn between existence, reality and time, in terms of their constituents. [...] In other words, *time cannot be external to existent reality but must be an existing part of it*.
>
> (Ingthorsson 2016, pp. 23–24, 29, [emphasis added], [9])

McTaggart says that if any reality is in time, then it must exist. That may be true, but it does not follow, that time itself is in time, that time is an existent somehow in the universe; the realm of concrete objects. Clearly, there are things in the universe that change, but time which is necessary for change, need itself not be in the universe and it need not change. There may be a content greater than the universe; there may be

a world that contains constituents that exist but are not in time. I shall argue that in such a world temporal *relations* and temporal *facts* (and all other relations and facts) exist. These relations and facts do not change but are what make change possible. Thus, it is a mistake to argue, as McTaggart does, that since the relations between events in a time-series do not change and the fact that say, an apple is green before it is red does not change, there is no time or change, unless A-change or becoming is introduced. To see what is involved in these points, it will be useful to consider a distinction that Ingthorsson blurs, between the notion of "fact" in McTaggart's substance ontology and in Russell's fact ontology.

On a substance ontology, the only ontological categories are substances, qualities and relations. Facts do not belong to a category of their own, and so must be understood as being either a substance, quality or relation. This is McTaggart's view, according to Ingthorsson:

> We need to become familiar with McTaggart's understanding not only of the general nature of the fundamental building blocks (substances, qualities, relations), but also of whatever has to do with the distinction and connection between thought and reality, such as *beliefs*, *assumptions*, *assertions*, *facts*, *truth* and *falsity*. With respect to the latter, one must understand that *they do not form a separate category of entities distinct from substances, qualities, and relations*.
>
> (Ingthorsson 2016, p. 16 [emphasis added], [9])

On the other hand, Ingthorsson also claims that

> *McTaggart's account of facts is in fact very similar to the account given by the logical atomists of an atomic fact*, notably that it consisted "either in the possession by a particular of a characteristic, or in a relation holding between two or more particulars" (Urmson 1960: 17).
>
> (Ingthorsson 2016, p. 28, [emphasis added], [9])

Surely, these two notions of "fact" although verbally similar—both are called "states of affairs"— are ontologically quite different.

In a fact ontology, such as logical atomism, there are particulars, non-relational and relational universals, and facts. If a particular exemplifies a non-relational quality, those two different *kinds* combine to form a fact

which is an entity over and above its constituents. It is a unity with complexity. If some particular stands in a relation to another, for example, if *a* is earlier than *b*, there is a temporal relational fact. Facts do not exist in time (or space), and the relational universals that are constituents in them do not exist in time (or space) either, although the terms (particulars) of temporal relations do exist in time.

The philosophy of logical atomism and the notion of an atomic fact is associated with Russell. He treats facts as a category of their own over and above their constituents when he says,

> Facts are [...] plainly something you have to take account of if you are going to give a complete account of the world. You cannot do that by merely enumerating the particular things in it: you must also mention the relations of these things, and their properties, and so forth, all of which are facts, so that facts certainly belong to an account of the objective world. ... and the things and their qualities or relations are quite clearly in some sense or other components of the facts that have those qualities or relations.
>
> (Russell 1918, pp. 191–192, [27])

A Russellian fact is a single (unity), yet complex entity that cannot be reduced to its constituents, and for that reason forms a separate category of entities distinct from substances (particulars), qualities and relations.

Despite Ingthorsson's attributing a substance ontology to McTaggart and his subsequent denial of the separate category of facts, McTaggart seems to recognize the category of fact and, like Russell, construes (some) facts as *timeless* even before his denial of the reality of time, since he says:

> When the subjective belief is eliminated, it seems to me that the truth goes with it and that we find ourselves left, not with a timeless, non-existent, and true proposition, but with nothing but the fact, which is not true (though it determines the truth of beliefs), which may or may not be *timeless*, and which, as we have seen above, is always in one way or another, existent.
>
> (McTaggart 1921, sect. 38, pp. 35–36, [13])

In the context of discussing the question of whether there are entities that are real, but non-existent, McTaggart distinguishes substance, quality and relations categorically:

> Qualities and relations are very different from substances, and the fact that a substance cannot be both existent and non-existent does not prove that qualities and relations—which are *universal*, and not *particular* as substances are—could not be existent in one aspect, and non-existent in another. (McTaggart 1921, sect. 5, p. 6, [13])

Although McTaggart ultimately rejects the reality of qualities and relations as real, but non-existent (unexemplified) he accepts Russell's distinction between universals and particulars. McTaggart claims "I am aware of an object or am acquainted with an object—the phrases are used synonymous—when 'I have a direct cognitive relation to that object'" (McTaggart 1921, sect. 44, p. 40, [13]). He then approvingly quotes Russell:

> When we ask what are the kinds of objects with which we are acquainted, the first and most obvious example is *sense-data*. When I see a colour or hear a noise, I have direct acquaintance with the colour or the noise. [...] But in addition to the awareness of the above kind of objects, which may be called awareness of *particulars*, we have also [...] what may be called awareness of *universals*. And universal *relations*, too, are objects of awareness; up and down, *before and after*, resemblance, and so on, would seem to be all of them objects of which we can be aware.
>
> (Russell 1912, pp. 209–212, [emphasis added]. In: McTaggart 1921, sect. 44, pp. 40–41, [13])

This is striking because Russell used direct acquaintance with *before and after* as grounds for maintaining that they are primitive, simple temporal relations.

Further evidence that McTaggart takes qualities and relations to be ultimate, and indefinable, ineliminable universals and distinct from particulars is his appeal to Russell's argument against reducing the common quality of two particulars to the simple relation of *exact likeness*

between them. McTaggart (1921, sect. 83, fn. 1, p. 83, [13]), cites Russell's argument in "On the Relations of Universals and Particulars," to justify the irreducibility of qualities as universals because the argument against qualities requires relations as (timeless) universals:

> Likeness at least, therefore, must be admitted as a universal, and, having admitted one universal we have no longer any reason to reject others. Thus, the whole complicated theory, which had no motive except to avoid universals, falls to the ground. [...] There must be relations which are universals in the sense that (a) they are concepts, not percepts; (b) *they do not exist in time*; (c) they are verbs, not substantives.
>
> (Russell 1911–12, p. 9, [24])

So, we can say that in certain respects McTaggart's ontology has similarities with Russell's. There exist substances (particulars) that stand in relations and have non-relational qualities. Qualities and relations are universals (they do not exist in space or time) and substances are particulars in that, if time exists, are the terms of temporal relations and so exist in time. Surprisingly, then, in (1921) Russell's influence on McTaggart is apparent in his treatment of relations, although there is also a fundamental difference.

McTaggart claims that the conception of relations is indefinable, since it is impossible to substitute for it any other concepts which can be taken as equivalent (McTaggart 1921, sect. 80, p. 80, [13]). He notes that theories have been put forth to demonstrate "that relation, though valid of existence, was not ultimate, but *definable* in terms of quality, so that statements about relations could be *translated* into statements about qualities" (McTaggart 1921, sect. 80, p. 81, [emphasis added], [13]). McTaggart maintains, however, that these theories are *mistaken*. "No fact which can be stated in terms of relations between substances can ever be stated in terms which omit the conception of relation" (McTaggart 1921, sect. 82, pp. 82–83, [13]). In other words, McTaggart *initially rejects* the view that relations can be reduced to or eliminated by the qualities of one or both of the terms of the relation.

Although McTaggart argues that relations are indefinable, irreducible and ineliminable in terms of qualities, he does say that if there are relations then there are qualities generated in each of the terms of the relation. McTaggart summarizes these points in the following passage:

> The conception of relation, then, must be accepted as valid of the existent. But it might be admitted to be valid, and yet denied to be ultimate and indefinable. It might be said that it really was true that substances were in relations, but that the fact expressed in this way could be expressed in terms of qualities only, without bringing in relations. But this also is false. [...] A relation may no doubt be based on a quality in each of its terms. But this does not mean that it can be *reduced* to those qualities. If A is larger than B, this relation may depend on the fact that A covers a square mile, and B covers an acre. [...] But a statement of the size of A and a statement of the size of B are not equivalent to a statement that A is larger than B, though the latter may be a certain and immediate conclusion from them.
>
> (McTaggart 1921, sect. 82, pp. 82–83, [13])

It is quite clear therefore that McTaggart (1921, [13]) accepts the existence of relations, including temporal relations, and in his discussion of cognition (1921, [13]), he quotes Russell who maintains that we are acquainted with them.

Thus, there are three views of relations that McTaggart considers. First, that relations are definable, reducible or eliminable in terms of qualities; a view McTaggart consistently rejects in (1921, [13]). Second, that relations are not definable or reducible to the qualities of their terms but depend on them; McTaggart's view in the passage just quoted. Third, Russell's view that relations are indefinable and irreducible to qualities and are not dependent on the qualities (temporal or otherwise) of their terms. McTaggart seems to have some sympathy with this Russellian view in his discussion of cognition quoted above. In the next section I shall argue that McTaggart's argument against the B-series, the A-series, and thus against the reality of time does not depend on the principle of temporal parity as Ingthorsson understands it, but on his changing analysis of temporal relations. For that reason, McTaggart's argument in (1927, [14]) should or at least could, as Broad maintained, be understood as a stand-alone argument that does not depend on his *a priori* metaphysics in (1921, [13]).

3 The R-theory and the Principle of Temporal Parity

Richard Gale refers to Russell as "the father of the B theory" and indeed, there is a good reason for him to do so, since McTaggart claims that "Mr. Russell's time-series [...] is identical with our B series [...] and the relation which unites the terms of the B-series is the relation of earlier and later" (McTaggart 1927, sect. 351, p. 31, [14]). However, the ontology of the B-series as McTaggart understands it is not the same as the Russellian time series as it should be understood. It will be useful, therefore to clarify Russell's authentic view in contrast to the B-theory which is McTaggart's misrepresentation of Russell but has nevertheless been accepted generally as Russellian.[3] A brief discussion of some of the differences between the Russellian theory of time ("R-theory" for short), and standard B-theoretic accounts as understood by defenders and critics alike is in order.

On typical interpretations, the ontology of B-time is construed as anti-realist because it denies that temporal passage is an objective, mind-independent feature of reality. For that reason, B-relations and the B-facts they enter, that alone constitute the foundation of the B-theory of time, are "nontransient" and static in that what appears to be the flow and flux of events in time —time's dynamism—is an illusion that would not exist without consciousness.[4] On the other hand, R-relations as given in experience are not static, but *dynamic*, and are the basis of our experience of transition and the passage of time.[5] Since the R-theory re-

[3]For a good discussion of McTaggart's misinterpretation of the Russellian view see Tegtmeier (2012, [32]).

[4]Not all B-theorists think of B-relations as static. See, for example, Savitt (2002, [30]), Deng (2013a, 2013b, [5–6]), Leininger (2014, [10]), and Mozersky (2015, [15]). For a critique of those who claim that the A-series, temporal passage and the dynamic aspect of time are illusions see Boccardi and Perelda (2017, [1]).

[5]Russell gives a colorful example of our experience of the earlier than relation in the following passage: "Immediate experience provides us with two time-relations among events: they may be simultaneous, or one may be earlier and the other later. These two are both part of the crude data; it is not the case that only the events are given, and their time-order, within certain limits, is a much given as the events. In any story of adventure you will find such passages as the following: 'With a cynical smile he pointed the revolver at the breast of the dauntless youth. "At the word *three* I shall fire," he said. The words one and two had already been spoken with a cool and deliberate distinctness. The word *three* forming on his lips. At this moments a blinding flash of lightning

jects the common view that B-time is a static, block universe, this last point deserves attention.

I have the knowledge that time *has passed* when, for example, I reflect on successive changes along my life's way (when I got married, when my first child was born, when my first grandchild was born and so on), or when I see that the position of the minute hand on my watch has changed its position. There is, however, a more immediate way in which we know that time *is passing*. This occurs when we are directly aware of passage, when we experience change in a single act of awareness. For example, if we look at a second hand of a watch or a flickering flame, we see the moving second hand at one place on the watch *before* the other, or one flicker of the flame occur *before* another in a single act of awareness. These are cases where we experience two stages of the second hand or flame occur in succession, one earlier than the other, and in so doing we are directly aware of a temporal transition or passage from one stage to the other. Similarly, when we hear the *successive* notes of a tune or feel the *successive* taps of a physician on our abdomen during a physical, we directly experience that time *is passing*.

On the R-theory, an appeal to the direct experience of succession in a single act of awareness is the basis for grounding our most basic experience of the flow or passage of time on mind-independent temporal earlier/later than *relations* alone; a view I will call the temporal relational theory or Russellian theory. On the R-theory, the commonsense belief that time passes is to be understood in terms of the *relation* of succession between earlier and later temporal objects. Thus, the passage of time consists in the succession of events throughout the history of the universe; one event – the earlier – being followed by another – the later.

Returning to the differences between B- and R-theorists, on standard (reductionist) B-theories, B-relations are analyzable in terms of causal relations whereas the R-theory takes R-relations as primitive and unanalyzable, relational universals that can be directly experienced. Russellian temporal relations are *external* relations, since "there are such facts as that one object has a certain relation to another, and that such facts *cannot* be reduced or inferred from, a fact about the one object only to-

rent the air.' Here we have simultaneity—not due, as Kant would have us believe, to the subjective mental apparatus of the dauntless youth, but given as objectively as the revolver and the lightening. And It is equally given in immediate experience that the words *one* and *two* come earlier than the flash'." (Russell, 1914, pp. 116–117, [25]).

gether with a fact about the other object only: they do not imply that the two objects have any complexity, or any *intrinsic* property distinguishing them from two objects which do not have the relation in question" (Russell 1966, pp. 139–40, [28]). For that reason, R-relations are neither analyzable in terms of A-properties of their terms nor do they depend on A-properties. Indeed, on the R-theory there are no such properties.

A last difference is particularly important for the discussion to follow. The B-theory is often identified with McTaggart's (1927, [14]) misinterpretation of Russell, according to which B-relations are unchanging and B-facts are *permanent* in that if *a* is *ever* (at any time) earlier than b, then *a* is *always* earlier than *b*. In contrast, R-theorists do not believe that either R-relations or R-facts exist in time, much less at every time, as McTaggart's interpretation implies. *Earlier than* is a *timeless* yet *dynamic temporal* relation. It is timeless because it does not exist in time; as a term of a temporal relation. It is dynamic because it is the ground of our experience of the passage of time; of *successively* existing temporal objects that exist tenselessly, that is, without tensed or A-properties. Similarly, time, understood as a Russellian series composed of a conjunction of R-facts, is timeless or atemporal. This view gives some meaning to an aphorism I favor, namely, *time is timeless*, or eternal in just this sense: though time contains temporal relations, time does not exemplify them.

Ingthorsson argues that McTaggart's paradox rests on a premise that can also be used to undermine the B-theory of time. The only view left standing is presentism. The premise is the "temporal parity thesis," the view that all times exist equally or co-exist, in a sense that is compatible with their being successive and not simultaneous or timeless. I disagree and shall argue that McTaggart's argument rests on a premise other than temporal parity whose implications refutes the B-series (as McTaggart understands it), the A-series (whose terms are past, present and future) and the A-theory including presentism, but when the dust settles, leaves the R-theory unscathed. This is a surprising conclusion, but true nonetheless, or so I shall argue.

The premise that plays a crucial role in McTaggart's argument against the B-series, the A-series and for the unreality of time is *not* the principle of temporal parity as Ingthorsson understands it, but McTaggart's analysis of "earlier than." He states that analysis in the following passage:

> The series of past, present and future is what we have called the *A* series, on which the *B* series of earlier and later is dependent. *The term P is earlier than the term Q, if it is ever past while Q is present, or present while Q is future.*
> (McTaggart 1927, sect. 610, p. 271, [emphasis added], [14])

This premise is nowhere argued for but is assumed in his argument against the B-series and the A-series. Moreover, it is incompatible with his account of relations in (1921, [13]), and so incompatible with the *a priori* metaphysics found there. Recall, that in (1921, [13]) he argued that relations are indefinable, and that while the terms of relations have qualities on which they may be based, they *cannot be reduced* to those qualities. However, in a footnote accompanying the passage above, McTaggart is claiming that "earlier than" *can be defined* in terms of temporal qualities (or relations to a term *X* outside the A-series). I shall argue that this analysis is crucial to his argument for the unreality of time but is incompatible with the temporal parity thesis, and his earlier view of relations. For that reason, McTaggart's (1927, [14]) argument can be construed as a stand-alone argument not dependent on his *a priori* metaphysics in (1921, [13]).

We shall see, even more importantly, McTaggart defends his analysis of *earlier than* by appealing to his rejection of the B-series, but his rejection of the B-series implicitly appeals to his analysis, and so massively begs the question against an R-theoretic interpretation of the B-series. It also implies an understanding of the temporal parity thesis that underscores a difference between McTaggart and R-theorists, even if McTaggart and B-theorists can be grouped together. To see what is involved in these points, let us return to McTaggart's text.

One problem with McTaggart's account of "earlier than" is that the word "while" implies that each of the disjuncts exist *in time* and thus the analysis is circular. For if *P* is past at $t1$, and *Q* is present at $t1$, then that is reducing "is past at" and "is present at" to the relations "is earlier than" and "is simultaneous with." If he denies absolute time, then "while" would imply simultaneity. Then Socrates is past is just as real as or *is simultaneous with* Oaklander is present. In that case, temporal parity implies co-existence in the sense of simultaneity, and that is incompatible with co-existing terms being successive. Finally, if "while" means co-present or existing at the same NOW, then we still haven't grounded

for example, Socrates existing earlier than Oaklander because Socrates being past, and Oaklander being present are both facts that exist now.[6]

Leaving that problem aside, McTaggart says that there seems to be a counterexample to his account of "earlier than" since in a durational present we are acquainted with the *earlier than* relation where both terms are *present* and thus are neither past or future. His reply is instructive:

> Two terms may both be present together, although one is earlier than the other. This is due to the fact that the present is a duration, and not an indivisible point. But the statement in the text remains an adequate *definition* of "earlier than," for although P and Q may *at one time* be in the same present, yet, *before* that, P is present while Q is future and *after* that, P is past while Q is present.
>
> (McTaggart 1927, sect. 610, fn. 1, p. 271, [emphasis added], [14])

McTaggart's response to the alleged counterexample is that if it is *ever* the case that P is past while Q is present, or P is present while Q is future," then P is *earlier than* Q, and the antecedent is satisfied because *before* P and Q are both present, P is present *while* Q is future, and *after* they are both present P is past *while* Q is present. Thus, McTaggart says that the statement in the text remains an adequate *definition* of "earlier than." However, to avoid an objection to his definition of "earlier than" by appealing to "before" and "after" is obviously circular. We saw that his appeal to "while" in his statement of the analysis of 'earlier than' is also problematic. This is important because it shows the need for primitive R-relations to account for the phenomenology and ontology of the *earlier than* relation and temporal passage. Indeed, McTaggart paradox results if we attempt to ground time without them, as we shall see.

McTaggart raises another objection to his analysis of "earlier than" that mirrors Russell's analysis of the tenses in "Our Experience of Time" (1915, [26]), that McTaggart was obviously familiar with. McTaggart says that since we can perceive the *earlier than* relation without perceiving the A-properties of *pastness* and *futurity*, perhaps we can take *earlier*

[6]Hope Sample has suggested to me that "while" could be understood in an atemporal sense of co-exist, but then co-existence would imply that each of the terms in each disjunct exist "eternally" and not successively, raising the specter of McTaggart's paradox.

than as primitive (as the R-theorist would have it) and define the future as what is later than the present, and the past as what is earlier than the present. He says:

> Since the present comprises different terms, of which any one will be earlier or later than any other, it might be thought that the fact that P was earlier than Q would be perceived when they were both present, and that "earlier than" *need not be defined in terms of the A series*. After this, it might be thought, the future may be defined as what is later than the present, and the past as what is earlier than the present. *Thus, the A series would be defined in terms of the B series, instead of the B series in terms of the A series*.
>
> (McTaggart 1927, sect. 610, fn. 1, p. 271, [emphasis added], [14])

In other words, McTaggart considers the view that the B-series alone constitutes time and that the A-series is not needed in a complete ontology of time given that we are directly acquainted with *earlier than* without being acquainted with A-properties.

McTaggart's response is familiar since it harkens back to his argument against the B-series in his chapter on "Time" (1927, [14]). He replies to the above argument by saying that to suppose that the B-series alone constitutes time

> [...] would be a mistake. For the series of earlier and later is a time series. We cannot have time without change, and the only possible change is from future to present, and from present to past. Thus, until the terms are taken as passing from future to present, and from present to past, they cannot be taken as in time or as earlier and later; and not only the conception of presentness, but those of pastness and futurity, must be reached before the conceptions of earlier and later, and not *vice versa*.
>
> (McTaggart 1927, sect. 610, fn. 1, p. 271, [emphasis added], [14])

McTaggart's argument against a B- or R-theoretic reduction of A-determinations is that earlier and later are temporal relations that generate a

time-series, only if their terms form an A-series and change A-properties with the *passage of time*. His justification is the argument quoted above that nothing changes in the B-series alone, since there is no temporal passage, and without events changing their A-determinations there are no temporal relations. I shall show, however, that this does not constitute an argument against the R-theory, since it mischaracterizes the R-series as a B-series and construes the B-series in a way that assumes McTaggart's definition of "earlier than." McTaggart's argument against the B-series also shows that temporal parity takes the terms of the B-series to be simultaneous, sempiternal or timeless, and not successive, and so cannot be employed against the R-theory. To see what is involved consider McTaggart's argument against the view that there *can* be a temporal series without the A-series. McTaggart argues that since the *relations* between the terms of the B-series are *permanent*, nothing changes on the B-series by coming into and going out of existence:

> If *N* is ever earlier than *O* and later than *M*, it will always be, and has always been, earlier than *O* and later than *M since* the relations of earlier and later are permanent. *N* will always be in the *B* series. And as, by our own hypothesis, a *B* series by itself constitutes time, *N* will always have a position in a time-series, and always has had one. That is, it always has been an event, and always will be one, and cannot begin or cease to be an event.
>
> (McTaggart 1927, sect. 310, p. 12, [14])

Generally, McTaggart says what he means, and means what he says. He does not say that "if it is ever *true* that *N* is earlier than *O* and later than M, it always will be *true* and always has been *true* that [...]." He is not talking about beliefs, but about facts, about time itself. So, the principle, if it is *ever* that case that P, then it is *always* the case that P, is stating a truth about the universe (the existent) that requires an ontological ground that *always* exists.[7]

[7] Although Ingthorsson would disagree with my literal interpretation of this passage (see Ingthorsson, 2016, p. 39, [9]), he does say "McTaggart consistently stays in object-language mode, assuming that he is talking about the world, and only ever takes a step back to talk about our talk of the world when he thinks that his particular use of words invites the risk of misunderstanding" (Ingthorsson 2016, pp. 89–90, [9]), and that would support my interpretation.

McTaggart's argument against the B-series and for the A-series (whether construed as terms having A-properties or standing in A-relations to a term outside the series) misinterprets the R-theory and assumes the existence of the A-series. McTaggart's argument misinterprets Russell since on the R-theory, temporal relational facts are *timeless* in the sense that they do not exist in time, i.e., they do not occupy moments (since there are none), they do not exemplify non-relational temporal properties (since there are none), and they do not stand in temporal relations (since relations of the first order can only have particulars as terms). On McTaggart's analysis of the B-series, however, B-relations between terms are *permanent*, they *always have been* and *always will be*. In other words, B-facts and the terms of those facts *always* exist; they exist at every time. "If *N* is *ever* earlier than *O* and later than *M*, it *will always be*, and *always has been*, earlier than *O* and later than *M*" (McTaggart 1927, sect. 310, p. 12, [emphasis added], [14]). This contradicts the notion of R-relations and R-facts, since if R-relations and R-facts do not exist in time then they cannot be permanent, and therefore cannot exist at every time. Why, then, in an argument against Russell would McTaggart assume that the B-series is permanent or *always* exists?

It is at this point that McTaggart's assumption of his definition of "earlier than" comes into play. If B-relations are analyzed in terms of different substances/events having incompatible A-properties at the same time (P is present *while* Q is future) or (P is past *while* Q is present), and so on, then the B-series would exist at every time. The B-series would be a permanent fact whose terms exists *at every time* with different temporal A-properties. Thus, without A-properties the series of the terms of the B-series would not be successive but either a simultaneous, sempiternal block, or timeless. Admittedly, on such an interpretation of the B-series nothing would change by coming into and going out of existence or by acquiring and losing a property. However, in an argument against R-relations, that are primitive, unanalyzable and indefinable relations, it obviously misunderstands R-relations (or the B-series as a Russellian would understand it) and begs the question.

McTaggart's argument that the characteristics of pastness, presentness and futurity must be *relations* to a term outside the time series and not qualities supports my reading of temporal parity as excluding coexisting yet successive times:

> Let us first examine the supposition that they are relations. In that case only one term of each relation can be an event or a moment. The other term must be something outside the time-series. For the relations of the A series are changing relations, and the relation of terms of the time-series to one another do not change. Two events are exactly in the same places in the time-series, relatively to one another, a million years *before* they take place, *while* each of them is taking place, and when they are a million years in the past. The same is true of the relation of moments to each other. Again, if the moments of time are to be distinguished as separate realities from the events which happen in them, the relation between an event and a moment is *unvarying*. Each event is in the same moment in the future, in the present, and in the past.
>
> (McTaggart 1908, p. 467, [emphasis added], [13])

Thus, for McTaggart, to say that a is earlier than b does not change means that they are in the same position in the time series *before* they take place, *after* they take place and when a and b are happening. However, if B-series facts exist at every time, then their constituents, a and b, always exist, and so are simultaneous or sempiternal, and not successive. Thus, NOW a is earlier than b, and it always has been the case and always will be the case that a is earlier than b. This is the ground of B-series facts always existing, but obviously, that cannot be an argument against the existence of the R-series without the A-series, unless it assumes an analysis of the B-series in terms of the A-series and confuses the R-series with the B-series.

Interestingly, some of the textual evidence that Ingthorsson gives to support his attributing the temporal parity principle to McTaggart implies, it seems to me, that past, present and future events do all co-exist, but simultaneously, and not successively. McTaggart says, "*Now* tomorrows weather is existent, for existence is as much a predicate of the future and past as of the present" (McTaggart 1921, sect. 6, fn. 1, p. 7, [13]). I think it is important to note that this quote implies not only that "Now tomorrows weather is existent," but also that "Now yesterday's weather is existent," and "Now, today's weather is existent." In other words, McTaggart's assertion of temporal parity—that past, present and future exist equally or co-exist—*does* imply that all the tensed facts in a

single A-series are *now,* hence simultaneous, and *does not* support that they exist successively.

McTaggart is not claiming that a single event is past present and future simultaneously, but that say, Socrates is past, Oaklander is present, and the 100th president of the US is future, are all contents in a single A-series, and thus all exist NOW, at the same time or simultaneously. Thus, *there is no ground for the different terms of a single A-series being successive even though they are equally real or co-existent*. For McTaggart (1927, [14]) temporal parity does not allow for all times to be co-existent and still be successive. McTaggart's notion of co-existence as applied to the terms of the A-series is simultaneity or timelessness, and therefore the temporal parity thesis is not something that McTaggart and R-theorists have in common. Thus, it is a mistake to claim that for the R-theorist "Socrates is in reality just as existent and real as we are now," (Ingthorsson 2016, p. 83, [9]) since for the R-theorist, Socrates is past, and Oaklander is present, are not facts that exist now because they do not exist in time at all.

By treating McTaggart and B-theorists on a par in accepting the principle of temporal parity, Ingthorsson fails to see that there is a fundamental difference between McTaggart and R-theorists. For McTaggart two objects/facts in *an* A-series cannot co-exist unless they are both NOW, and so exist at the same time. The R-theorist need not accept temporal parity in that sense. The R-theorist does not accept that Socrates is just as real and existent as we are now, that there is a sense in which Socrates "still exists."

It might be objected that if R-relations are atemporal, then that is one notion of "permanence" and in this context to be permanent means to exist independently of any change. Thus, McTaggart could be making the point that the B-series does not involve change since the relations that generate it do not change. *But it is a mistake to suppose that the basis of change must itself change; that the ground of temporality in the world must itself be temporal in the sense of existence in time*.[8]

[8]For a fuller discussion of this point including a reply to the "No change" objection to the R-theory in the context of Ingthorsson's critique of the B-theory, see my review of Ingthorsson's book (Oaklander, forthcoming 2019, [23]). For an account of emerging and passing away consistent with the R-theory see Tegtmeier (1999, [31]).

4 McTaggart's Paradox

To begin our discussion of McTaggart's paradox[9] let me call your attention to an ambiguity, already alluded to, in the notion of the A-series. The A-series may be a single A_1-series, or the A-series can be an A_2-series of A_1-series. Consider the following quote where McTaggart characterizes an A_1-series:

> We must begin with the A series, rather than with past, present, and future, as separate terms. And we must say that a series is an A series when each of its terms has, to an entity X outside the series, *one, and only one*, of three indefinable relations, pastness, presentness, and futurity, which are such that all the terms which have the relation of presentness to X fall between all the terms which have the relation of pastness to X, on the one hand, and all the terms which have the relation of futurity to X, on the other hand.
> (McTaggart 1927, sect. 328, p. 20, [emphasis added], [14])

A single A-series has terms that each co-exist, and Ingthorsson claims that co-existence is compatible with their existing in succession, but that is incorrect since each term in a single A-series is NOW either past, present, or future. It is correct that a, b, c exist equally, but not that they exist in succession.

There are two reasons why a single A_1-series is not a genuine temporal series whose terms exist in succession. First, because there is no change in a single A-series because there is no term that has a property and then loses it. Second, although his definition of "earlier than" would suggest that a single A-series whose terms have different A-properties is a temporal series, that is not in fact the case for if P is past *while* Q is present, then P and Q exist at the same time either *simultaneously*, if in B-time, presently, in A-time, or at t_1, if time is absolute. Thus, there is no ground for *P is earlier than Q* in a single A-series, for to generate a temporal series that contains "real change" there must be *temporal passage* and that requires a series of A-series. An A_2-series of A_1-series in which each different A_1-series have terms with incompatible A-properties. For example, in one A_{1a}-series, e is present and e' is future, and in another

[9] For my earlier account of McTaggart's Paradox see Oaklander (2002, [16]).

A_{1b}-series e is past and e' is present. Of course, an A_2-series of such a conjunction of A_1 series is not yet a *temporal* series for unless we introduce time in some way the A_2-series is contradictory. What, then, is to be done? McTaggart's answer is startling:

> When we say that the B series is a series of changes, we do not, of course, mean that the terms change their places in the series. If one term is ever earlier than another, it is always earlier than that other. But the B series is a time-series, and time involves change. And the change in the terms of the B series is that they are *successively* present (*passing* from futurity to presentness, and from presentness to pastness). It is first an *earlier* term which is present, and then a *later* one. (McTaggart 1927, sect. 698, p. 347, [emphasis added], [14])

What is key here is that the ground of the passage of time from the future to the present is the relation of *succession*. The A_2 series of A_1-series is a temporal series because the generating relation is *earlier than* since "it is first an earlier term [in an A_{1a} -series] which is present, and then a later term (in an A_{1b}-series) that is present." *In other words, A-theoretic change or temporal passage presupposes succession*.[10] It is not sufficient for change that one and the same event/object/time have incompatible temporal properties, they must have them successively. However, the appeal to succession to ground the passage of time gives rise to either the unreality of time, due to a vicious infinite regress if succession is analyzed A-theoretically in terms of McTaggart's definition of "earlier than," or the R-theory of time if succession is analyzed R-theoretically, since then R-relations are more fundamental than A-properties, and temporal passage and temporal becoming are grounded in R-relations.

To put this point slightly differently, let's go back to his definition: "The term P is earlier than the term Q, if it is ever past while Q is present, or present while Q is future" (1927, sect. 610, p. 271, [14]). Since neither disjunct alone could *ground P* is earlier than *Q* (even if each could entail it), the disjunction should be a conjunction. In that case, P is past *while* Q is present, *and* P is present *while* Q is future. If "while" is non-temporal,

[10] This point is also argued for in Boccardi [2].

then a contradiction exists since P and Q would timelessly have incompatible properties. On the other hand, if "while" is temporal, then it is assuming time and so we must ask, what more is needed to give us time and change? McTaggart answers that question, and his answer is quite revealing. To repeat:

> And the change in their terms of the B series is that they are **successively** present (passing from futurity to presentness, and from presentness to pastness). It is *first* an **earlier** term which is present, and *then* a **later** one.
>
> (McTaggart 1927, sect. 698, p. 347, [14])

In other words, the ground of change in the B-series is the transition from one present to another; the transition from one A_{1a}-series to another A_{1b} -series is that the terms of the A_2-series are **successively** present! However, if *succession* is a primitive R-relation and the basis of temporal passage or temporal transition then McTaggart's definition of "earlier than" in terms of A-properties is undermined. On the other hand, if *succession* from one present to another is analyzed in terms of his definition of "earlier than," that would give rise to another third level A_3-series whose terms are a series of A_2-series, with incompatible A-properties, that is contradictory and static without introducing succession and thus leading to a vicious infinite regress. Thus, without R-relations and the R-series as the basis of passage, the result is a contradiction or a vicious infinite regress, but with R-relations, the A-series is not needed for time to pass. For that reason, the significance of McTaggart's argument is not that time is unreal, but that time requires temporal passage which can only be grounded by appealing to a primitive R-relation of succession.

McTaggart's argument can also be applied to presentism. For suppose we consider his definition of "earlier than" and say that only one of the disjuncts, that is, only one A-series exists, the one that is *now*. Since it is the case that say, P is past *while* Q is present, we can then maintain that "while" designates the *present* time. The seemingly advantage of presentism in this case is that all tensed judgments, including those about the future and the past for example, "It will rain" and "It did snow" and "It is now sunny," are all grounded in the present. The single A-series that is present. Moreover, given his definition of "earlier than," a single

A-series presumably also grounds the existence of that relation. Importantly, there is no contradiction since no single term of an A-series has incompatible A-properties. These are all advantages of a McTaggartian conception of presentism.

Nevertheless, these apparent advantages for presentism come to naught since a single (present) A-series is not sufficient to constitute time. To have time there must be *passage* and for that presentism must account for continual becoming or absolute becoming. Passage involves a *transition* from the existence of one single A-series to the existence of another single A-series. In other words, one present (a single A-series) must go out of existence and another come into existence. However, transition is from existence to non-existence/existence (existence of an A_{1a} -series to ceasing to exist of A_{1a} and coming to exist of A_{1b}), and that involves succession.

In other words, a single A-series presupposes a series of A-series since, as Richard Gale has noted:

> It can easily be shown that if there is one A-series there must be a series of A-series. Assume that the A-series consist of events M, N and O, which are respectively past, present and future. A past (future) event by definition is one which was (will be) present. Thus, if there is one A-series there is becoming—a series of A-series; and if the A-series is objective then so too is becoming. (Gale, 1969, p. 190, [7])

Note, however, that McTaggart's point is that a sequence of A-series does not constitute a temporal sequence or series unless the generating relation is a temporal relation. Thus, if $P(Q)$ is ever past (present) then at an *earlier* time $P(Q)$ had to be present (future). In other words, a single A-series is not itself a temporal series since it does not involve change. What is needed to generate time is passage, but passage requires a succession of present A_1-series or at least a succession of present times for different A_1 series to become present at. Thus, without a primitive notion of *earlier than* a single A-series whose terms are past, present and future, could not exist. Of course, with a primitive *earlier than* relation there would not be an A-series either since a primitive *earlier than* relation is nothing other than the R-relation which therefore must exist, and an R-relation being an external relation does not have terms with A-properties.

Clearly, Ingthorsson would reject this McTaggartian understanding of presentism because he rejects A-properties (past and future events), and "denies that tensed passage occurs at all" (Ingthorsson 2016, p. 128, [9]). Nevertheless, I think the line of argument I have developed in interpreting McTaggart applies to his form of presentism as well. For although Ingthorsson rejects *tensed* passage, he undoubtedly accepts *temporal* passage since he believes that the ordinary conception of change is dependent on it. He says,

> Change is a difference in the properties of an object that remains numerically identical through the change, i.e., 'genuine change'. This conception of change requires that things persists by *enduring*, i.e., that they come to exist at many times by *passing* as numerically identical three-dimensional bodies through time. [...] It is the view that ordinary material objects are three-dimensional *things that move as numerical wholes through a succession of times*.
>
> (Ingthorsson 2016, p. 94, [emphasis added], [9])

Ingthorsson highlights the need for temporal passage in endurantism when he says: "It is indeed some form of *temporal passage* that is supposed to allow three-dimensional objects to come to exist at many times and yet exist *completely* and *only* at each of those times., i.e. without having parts 'lying around' at other times" (Ingthorsson 2016, p. 95, [emphasis added], [9]). Ingthorsson makes it clear that temporal passage is "what allows three-dimensional particulars to be at many times in succession" (Ingthorsson 2016, p. 99, [9]) and to be wholly present at those times. His form of presentism involves "permanent" material substances that do not come into being or go out of being, but "'remains' continuously in the present" (Ingthorsson 2016, p. 138, [9]) as they pass through the flow of time and change.

There are several problems with this analysis that lead us once again, it seems to me, to the R-theory. Ingthorsson maintains that just as an object can move from one place to another only if it passes as a whole from one place to another, an object can pass from one moment to another only if it is wholly contained in each moment through which it passes (Ingthorsson 2016, p. 99, [9]). The first point I want to make is that the analogy assumes the existence of times that exist before and after a substance arrives at them, just as spatial places must exist before and after

an object leaves and arrives at them. The idea of motion assumes that the place where we left from still exists and the place that we are moving to already exists. So, if an object, x, leaves one time, arrives at another time, and moves toward another time, this implies that more than the present time exists. Thus, the notion of temporal passage, as Ingthorsson unwittingly conceives of it, involves the co-existing of times that are either not successive (if only the present exists), but are simultaneous, or they are successive but then it is not the case that only the present exists and is incompatible with his denial of cross-temporal relations.

Thus, Ingthorsson is faced with a dilemma. If the passage of objects through time is grounded in the succession of times through which objects move, then since succession is a cross-temporal relation and everything cannot be grounded in the present. Moreover, it treats the time-series as a series through which objects move leading to all the problems of the moving spotlight view, and further undermining presentism. Similarly, if a substance moves from one time to another time, then those times must exist (or permanently remain in existence) for substances to move toward and away from. If moments permanently remain in existence, however, then they are simultaneous and sempiternal and so are not successive, unless one accepts McTaggart's definition of "earlier than" with the subsequent vicious infinite regress to follow.

On the other hand, if past times no longer exist and future times do not yet exist, then the notion of passing from one time to another makes no sense, since there are no times or a temporal series through which an object moves. If a substance *comes to exist* at a certain time, *it is moving towards something* that already exists "waiting" for the substance to arrive. A substance cannot move to a time and thus *come to exist at that time* unless there is *something* it is moving toward. The best it can do is *come into existence*, through absolutely becoming, at a time that also comes into existence. That would preserve presentism, but it would do away with temporal passage because the notion of moving successively *from* one time *to* another has no ground if there does not exist a temporal series, that is, if only the present time exists.

Ingthorsson claims that if there is no passage, as on the B-theory, then everything remains permanently at its own temporal position. But that implies that without passage the world is a *totum simul*. The result is that there are no temporal relations. However, McTaggart's point is that temporal passage requires succession. Both the presentist and the

R-theorist reject temporal passage as tensed property change. The Russellian rejects it because there are primitive temporal relations whose terms do not exemplify A-properties. Ingthorsson rejects tensed properties and R-relations (cross-temporal relations) because the past and the future do not exist but puts nothing in their place. Therefore, succession has no ontological ground in his universe and therefore, the positions that his substances move through are not temporal positions or times, but spatial positions that are sempiternal or timeless positions. Certainly, they are not successive.

Change for Ingthorsson is "the objective loss and acquisition of properties by an enduring portion of matter. When something changes a qualitative state ceases to exist as another begins to exist, and never is there a cross-time relation between two qualitative states of the same substance existing at different times. [...] Change is "one state goes out of being while another begins to be" (Ingthorsson 2016, pp. 135–136, [9]). However, if the time at which a qualitative state of a substance wholly present comes to exist is simultaneous with the substances' arrival, and a qualitative state of that "same" substance ceases to exist at the same time as the substances' arrival, then it is not the same substance that loses and acquires a qualitative state. If a time comes into existence absolutely, then everything at that time must also come into existence absolutely. There is no temporal passage from one time to another, but just creation *ex nihilo* at that time. For if there is no passage, no succession of times though which substances move, then not only the qualitative state, but the substance having that qualitative state, goes out of existence and the substance that comes into existence with a qualitative state is not the same as the substance that ceased to exist with a different qualitative state. Thus, there is not a single substance that changes from one present to another.

Ingthorsson claims that,

> The alterations that take place are a question of qualitative states coming into and going out of being, and they are provoked by the influence exerted between different portions of matter. Here we find presentism thoroughly embedded in a metaphysics of material nature. And it is the most sparse ontology I know of. (Ingthorsson 2016, p. 138, [9])

Indeed it is, or seems to be, a naturalist ontology where everything ex-

ists in space and time. This is reinforced by his claim that "The core idea of presentism, as standardly formulated, is that the concretely existing present is meant to ground everything" (Ingthorsson 2016, p. 135, [9]). How, then, can there be a succession of times that substances pass through? Surely Ingthorsson cannot ground the succession of times in a concretely existing present unless he accepts McTaggart's definition of succession that we have seen, leads to a vicious infinite regress, that is, McTaggart's paradox. For the R-theorist, on the other hand, in order to "ground everything" we must recognize not only the concrete particular existent, but the abstract non-spatial and non-temporal realm. Thus, the dispute between presentism and the R-theory is a debate between the naturalist and the ontologist à la Grossmann.

5 Conclusion

My understanding of McTaggart's refutation of the A-theory including presentism, also enables us to understand the overarching point of McTaggart's main argument for the unreality of time. After arguing for the existence of A-change, he claims that the obvious reply to the attribution of incompatible A-properties/relations to events is that they have those properties *successively*. What, then, we must ask "does it mean to say that events or moments of time have A-properties *successively*?" His reply appeals to moments of time and A-properties (*has been*, *will be*, and *is now*) as he later claims in his definition of "earlier than" in section 610. Obviously, that does not work, as he will subsequently show, since the passage of time involves different events/times *successively* becoming present. Thus, if succession is analyzed in terms of a single A_1-series, or an A_2-series of A_1-series, or an A_3-series of A_2-series, the analysis either does not yield change, or yields a contradiction, that cannot be removed without appealing to succession and the subsequent infinite regress.

McTaggart summarizes his argument in the following passage:

> The attribution of the characteristics past, present, and future to the terms of any series leads to a contradiction, unless it is specified that they have them *successively*. *This means, as we have seen, that they have them in relation to terms specified as past, present, and future*. These again, to avoid a like con-

> tradiction, must in turn be specified as past, present, and future. And, since this continues infinitely, the first set of terms never escapes from contradiction at all.
> (McTaggart 1927, sect. 332, p. 22, [emphasis added], [14].)

We see then that his argument in 1927 does not depend on the temporal parity as Ingthorsson understands it, or the *a priori* metaphysics of 1921, but on his analysis of succession that is not argued for but assumed. In a footnote to the passage just quoted, McTaggart makes it clear that the vicious infinite regress arises from the attempt to avoid the contradiction in the attribution of incompatible A-characteristics to the terms in each A_1-series by appealing to succession and then analyzing succession in terms of incompatible A-characteristics. This passage is important enough to quote at length:

> It may be worthwhile to point out that the vicious infinite does not arise from the impossibility of *defining* past, present, and future, without using the terms in their own definitions. On the contrary, we have admitted these terms to be indefinable. It arises from the fact that the nature of the terms involves a contradiction, and that the attempt to remove the contradiction involves the employment of the terms and the generation of a similar contradiction.
> (McTaggart 1927, sect. 332, fn. 1, p. 22, [emphasis added], [14].)

To conclude, I believe to have shown three things. First, that McTaggart's argument against Russell begins by assuming his analysis of the "earlier than" relation as stated in 1927 section 610, and he uses that analysis in his main argument against the B-series. This has given rise to an A-theoretic misunderstanding that distorts the R-theory. Second, McTaggart's analysis of "earlier than" undermines all versions of the property and the presentist versions of the A-theory, as his argument for the unreality of the A-series sought to demonstrate. Third, since McTaggart is clear that the passage of time and the possibility of change requires *succession*, I conclude that McTaggart has shown that given his analysis of "earlier than," time and change are unreal. Therefore, to preserve the reality of time and change, we must reject McTaggart's analysis of "earlier than," recognize that it presupposes a *primitive* notion

of succession and assert that the relation that generates change is the simple, unanalyzable R-relation of *earlier than*.

Acknowledgements

I wish to thank Emiliano Boccardi, Silvano Miracchi, Hope Sample and Erwin Tegtmeier for their helpful comments on earlier drafts of this paper.

Bibliography

[1] Boccardi, E. and Perelada, F. (2017). The Delusive Illusion of Passage. *Analysis* **78**(3): 387–396.

[2] Boccardi, E. (2018). Turning the Tables on McTaggart. *Philosophy* **93**(3): 1–169.

[3] Broad, C.D. (1933). *Examination of McTaggart's Philosophy*, vol. I. London: Cambridge at the University Press.

[4] Broad, C.D. (1938). *Examination of McTaggart's Philosophy,* Vol. II. Pt I. London: Cambridge at the University Press.

[5] Deng, N. (2013a). Fine's McTaggart, Temporal Passage, and the A-versus B-Debate. *Ratio* **26**(1): 19–34.

[6] Deng, N. (2013b). Our experience of passage on the B-theory. *Erkenntnis* **78**(4): 713–726.

[7] Gale, R.M. (1969). *The Language of Time*. London: Routledge and Kegan Paul.

[8] Grossmann, R. (1992). *The Existence of the World: An Introduction to Ontology*. London: Routledge.

[9] Ingthorsson, R.D. (2016). *McTaggart's Paradox*. London: Routledge.

[10] Leininger, L. (2014). On Mellor and the Future Direction of Time. *Analysis Reviews* **74**(1): 148–157.

[11] Lowe, E. J. (1987). The Indexical Fallacy in McTaggart's Proof of the Unreality of Time. *Mind* **96**(381): 62–70.

[12] McTaggart, J.M.E. (1908). The Unreality of Time. *Mind* **24**(68): 457–474.

[13] McTaggart, J.M.E. (1921). *The Nature of Existence*, vol. I. *Cambridge*: Cambridge University Press.

[14] McTaggart, J.M.E. (1927). *The Nature of Existence*, vol. II. C.D. Broad (ed.). Cambridge: Cambridge University Press.
[15] Mozersky, M.J. (2015). *Time, Language, and Ontology: The World from the B-theoretic Perspective*. Oxford: Oxford University Press.
[16] Oaklander, L.N. (2002). McTaggart's Paradox Defended. *Metaphysica* **3** (1): 11–25. Reprinted in: Oaklander, L.N: *The Ontology of Time*. Amherst. NY: Prometheus Books, 2004: pp. 51–62.
[17] Oaklander, L.N. (2012). A-, B- and R-theories of Time: A Debate. In: Bardon, A. (ed.): *The Future of the Philosophy of Time*. New York: Routledge: pp. 1–24.
[18] Oaklander, L.N. (2014a). Dolev's Metaphysical Anti-Realism: A Critique. In L. N. Oaklander (ed.): *Debates in the Philosophy of Time*. New York: Bloomsbury Publishing: pp. 1–29.
[19] Oaklander, L.N. (2014b). Temporal Realism and the R-theory. In Bonino, G., Cumpa, J. and Jesson, G. (eds.): *Defending Realism: Ontological and Epistemological Investigations*. Boston: De Gruyter: pp. 122–139.
[20] Oaklander, L.N. (2015). Temporal Phenomena, Ontology and the R-theory. *Metaphysica* **16**(2): 253–269.
[21] Oaklander, L.N. (2016). Common Sense, Ontology and Time: A Critique of Lynne Rudder Baker's View of Temporal Reality. *Manuscrito – Rev. Int. Fil. Campinas* **39**(4): 117–156.
[22] Oaklander, L.N. (2017). The Best of Times. *The Philosophers' Magazine* **79**(4): 44–49.
[23] Oaklander, L.N. (forthcoming, 2019). Review of R. D. Ingthorsson, *McTaggart's Paradox*. *Metaphysica*.
[24] Russell, B. (1911–12). On the Relation of Universals and Particulars," Proceedings of the Aristotelian Society: pp. 1–24. In: Russell, B. (2001). *Logic and Knowledge*: *Essays 1901–1950*. Marsh, R. (ed.). London: Allen and Unwin, New York: Macmillan: pp. 115–124.
[25] Russell, B. (1914). *Our Knowledge of the External World*. London: Allen and Unwin.
[26] Russell, B. (1915). On the Experience of Time. *Monist* **25**: 212–33. In: Oaklander, L.N. (ed.): (2000). *The Philosophy of Time: Critical Concepts in Philosophy*, vol. I: pp. 174–187.
[27] Russell, B. (1918). Lectures on The Philosophy of Logical Atomism. *Monist*. In: Russell, B. (2001). *Logic and Knowledge*: *Essays 1901–1950*. Marsh, R. (ed.). London: Allen and Unwin, New York:

Macmillan: pp. 176–281.
[28] Russell, B. (1966). The Monistic Theory of Truth. In: *Philosophical Essays*. London: George Allen Unwin Ltd: pp. 131–146
[29] Russell, B. (2001). *Logic and Knowledge*: *Essays 1901–1950*. Marsh, R. (ed.). London: Allen and Unwin, New York: Macmillan.
[30] Savitt, S. (2002). On Absolute Becoming and the Myth of Passage. In Callender, C. (ed.): *Time, Reality, and Experience*. The Royal Institute of Philosophy Supplement 50. Cambridge: Cambridge University Press: pp. 153–167.
[31] Tegtmeier, E. (1999). Parmenides' Problem of Becoming and its Solution. *Logical Analysis and History of Philosophy* **2**: 51–65.
[32] Tegtmeier, E. (2012). Error: Temporal Change. *Revue Romaine de Philosophie* **56**: 89–96.
[33] Urmson, J.O. (1960). *Philosophical Analysis*. London: Oxford University Press.

Dummett on McTaggart's Proof of the Unreality of Time

Brian Garrett
Australian National University, Australia
Brian.Garrett@anu.edu.au

Abstract

Michael Dummett's paper "A Defence of McTaggart's Proof of the Unreality of Time" put forward an ingenious interpretation of McTaggart's famous proof. My aim in this discussion is not to assess the cogency of McTaggart's reasoning, but to criticise Dummett's interpretation of McTaggart.

Keywords: McTaggart's paradox, Michael Dummett, temporal immersion, observer-independence.

1 Introduction

The reasoning of McTaggart's 1908 article 'The Unreality of Time' runs as follows. We distinguish positions in time in two ways: a permanent B-series (in which events and facts are distinguished using the relations of *earlier than* and *later than*) and a dynamic A-series (in which events and facts are future, then present, then past). Both series are essential to time, yet the A-series is more fundamental since only it allows for

change. (McTaggart 1908, p. 458, [2]). This concludes the first part of McTaggart's reasoning: his argument for the fundamentality of the A-series.

Having established this conclusion, McTaggart then claims that the A-series "[...] involves a contradiction." (McTaggart 1908, p. 466, [2]). His argument for the contradiction is seemingly straightforward: past, present and future are "incompatible determinations" yet "[...] every event has them all." (McTaggart 1908, p. 469, [2]). This argument is typically known as McTaggart's Paradox. McTaggart is aware of a natural rejoinder to his argument. No event, it will be urged, is simultaneously past, present and future, only successively, and from this no contradiction follows. But, claims McTaggart, this rejoinder entails either a vicious circle or a vicious infinite regress, and so the contradiction is not removed.

In sum, McTaggart's first argument establishes that the A-series is fundamental to time. His paradox then establishes that the A-series is contradictory. From these conclusions it follows that time is unreal. My aim here is not to assess McTaggart's reasoning, but to call into question Michael Dummett's well-known interpretation of McTaggart's proof of the unreality of time. (Dummett 1960, [1]).

In his paper Dummett does not use the terms 'A-series' and 'B-series'. Instead, he talks of "facts of kind (a)" *viz*., facts into the statement of which tensed expressions enter essentially. (Dummett 1960, p. 500, [1]). Presumably, facts of kind (b) would be facts into the statement of which no tensed expressions enter (or do so inessentially). Clearly, facts of kind (a) - tensed facts - are meant to correspond to A-series facts (e.g., the fact that Hitler's death is past), and facts of kind (b) - tenseless facts - are meant to correspond to B-series facts (e.g., the fact that Hitler's death is later than Caesar's death).

2 McTaggart's argument

Having thus set things up, Dummett begins by making the following claim. With regard to McTaggart's arguments, "[p]art two depends upon part one [...]" (Dummett 1960, p 500, [1]). That is, the success of McTaggart's Paradox, a plausible rendering of which Dummett sketches in his opening pages, depends upon the success of McTaggart's argument for the fundamentality of the A-series. This might seem a puzzling

claim. Could not one hold that the A-series is contradictory even if it is not fundamental?

However, I take it that Dummett's point is that McTaggart's Paradox (part two) is directed against the A series as understood by the A-theorist. That is, McTaggart's Paradox is intended to show that the A-series is contradictory on the assumption that tensed facts are fundamental and (hence) irreducible. No one thinks that the A-series is contradictory if tensed facts are taken to be reducible to B-theoretic facts (e.g., it's not contradictory for X to be earlier than t_2 but later than t_1). McTaggart's Paradox is directed against the A-theory of time, i.e., the theory according to which the A-series is fundamental and consistent. So understood, Dummett's claim is correct: McTaggart's proof of the inconsistency of the A-series presupposes the fundamentality of that series. This explains why Dummett continues as follows:

> [...] it is because the analogue of part one does not hold for space or for personality that the analogue of part two for space or for personality has no force.
>
> (Dummett 1960, p. 500, [1])

The analogue of part one does indeed fail for space and personality. Spatially and personally token-reflexive expressions ('here', 'there', 'I', 'you', etc.,) need not feature in a full description of reality. A description using only spatial co-ordinates and personal proper names would suffice.

The conclusion of McTaggart's first argument - that the A-series is fundamental to time - implies the *tensed thesis* (as we can call it) that what is in time cannot be fully described without the use of tensed expressions. Dummett is sympathetic to this A-theoretic thesis and offers his own argument for it. Consider any description of events containing no tensed expressions. We can, he says, always ask the question "And which of these events is happening *now*?" This question, Dummett thinks, deserves an answer, yet can be given one only if tensed expressions are added to the description. Hence, the tensed thesis is true. (Dummett 1960, p. 591, [1]).

This argument is unconvincing. Unlike the A theorist, the B-theorist rejects the tensed thesis and accepts instead the *tenseless thesis* that what is in time can be fully described without the use of tensed expressions. According to the B-theorist, Dummett's question is either illegitimate or else can be answered in tenseless terms. If Dummett's question is

asked from 'outside time', it makes no sense, just as the question 'What is happening *here*?', asked from 'outside space', makes no sense. For the B-theorist, a question containing a temporal indexical can only be asked and answered from a position in time. In that case, the answer to Dummett's question ("And which of these events is happening *now*?") is straightforward: those events whose occurrence is simultaneous with the posing of that very question.

3 Temporal versus spatial immersion

Dummett takes the tensed thesis to be equivalent to the *temporal immersion* thesis that a description of events in time can only be "[...] given by someone who is himself in that time." (Dummett 1960, p. 501, [1]). Dummett is well-disposed towards this thesis, but unhesitatingly rejects the *spatial immersion* thesis that a description of objects in space can only be given by someone who is himself in that space. He writes:

> [...] the use of spatially token-reflexive expressions is not essential to the description of objects as being in a space. That is, I can describe an arrangement of objects in space although I do not myself have any position in that space. An example would be the space of my visual field. In that space there is no here or there, no near or far: I am not in that space. We can, I think, conceive, on the strength of this analogy, of a being who could perceive objects in our three-dimensional physical space although he occupied no position in that space. He would have no use for any spatially token-reflexive expressions in giving a description of the physical universe, and yet that description might be a perfectly correct description of the objects of the universe as arranged in space. (Dummett 1960, pp. 500-501, [1])

Dummett's reasoning in this passage is somewhat convoluted. He cites the space of one's visual field as a counterexample to the spatial immersion thesis. On the basis of the analogy with the space of one's visual field, we can imagine a being who could perceive objects in our three-dimensional physical space without occupying any position in

that space. Such a being could give a full description of objects in physical space without occupying any position in that space, thus yielding another counterexample to the spatial immersion thesis.

The spatial immersion thesis may well be implausible, but the analogy with the space of the visual field does not help Dummett's case. If the objects of my visual field are non-physical sense-data, they occupy no space. In that case, the 'space' of my visual field is irrelevant to the spatial immersion thesis. If, instead, the objects of my visual field are those objects in my immediate physical environment, then I am in the same space as them. Again, we have no counterexample to the spatial immersion thesis. (Thomson 2001, pp. 243-247, [3]).

4 Observer-independence

With regard to part two of McTaggart's argument, Dummett asks "[...] does not the objection we considered - that McTaggart's attempt to uncover a contradiction rested on a neglect of the obvious properties of token-reflexive expressions - at least invalidate part two of the argument?" (Dummett 1960, p. 501, [1]. Rightly or wrongly, Dummett uses 'token-reflexive' interchangeably with 'indexical'.) The objection Dummett is alluding to holds that, if McTaggart's argument for a contradiction in the A-series were sound, we could equally well argue for the inconsistency of space and personality by showing that every place can be both 'here' and 'there', and every person can be both 'I' and 'you'. Since the latter arguments are confused, so is McTaggart's.

It is odd that Dummett asks this question at this point since he already has the means to answer it, *viz*., by appeal to the falsity of the analogue of part one for space and personality. However, instead of giving this answer, Dummett takes a new tack and ascribes to McTaggart the assumption that:

> [...] reality must be something of which there exists in principle a complete description. I can make drawings of a rock from various angles, but if I am asked to say what the real shape of the rock is, I can give a description of it as in three-dimensional space which is independent of the angle from which it is looked at. The description of what is really there, as it really is, must be independent of any particular point

> of view. Now if time were real, then since what is temporal cannot be completely described without the use of token-reflexive expressions, there would be no such thing as the complete description of reality.
>
> (Dummett 1960, p. 503, [1])

According to Dummett's use of 'complete', a complete description is "independent of any particular point of view", i.e., observer-independent. Dummett thus ascribes to McTaggart the *observer-independence* thesis that there can be an observer-independent description of temporal reality.

As the above quote makes clear, Dummett takes the tensed thesis to imply the falsity of the observer-independence thesis. If what is temporal cannot be completely described without the use of perspectival terms such as 'past', 'present' and 'future', then what is temporal can only be fully described from the perspective of a being in time. This seems correct. Dummett takes McTaggart to endorse both the tensed and observer-independence theses, and hence to be (knowingly) in the grip of a contradiction which can be avoided only by declaring time to be unreal. Hence, McTaggart is forced to his infamous conclusion that time is unreal.

However, Dummett offers no textual support for this interpretation of McTaggart. No passage is cited in which McTaggart endorses the observer-independence thesis. It is, in addition, a strange and self-stultifying argument. Any supporter of the tensed thesis will automatically reject the observer-independence thesis, and *vice-versa*. Moreover, the observer-independence thesis has no theory-independent plausibility. Its acceptance could only be motivated by a belief in the tenseless thesis (i.e., the denial of the tensed thesis). For these reasons, then, Dummett has given us no reason to revise the standard, and textually-grounded, interpretation of McTaggart's argument which emphasises the role played by McTaggart's well-known (if ill-understood) proof of the self-contradictory nature of the A-series, as that series is understood by the A-theorist.

Furthermore, the argument which Dummett himself endorses – (*contra* McTaggart) time is real; the tensed thesis is true; so observer-independence is false – is an argument that any A-theorist should endorse. Evidence that this is Dummett's argument can be seen from

the following remark: "If this last piece of reasoning, to the effect that the belief that time is unreal is self-refuting, is correct, then McTaggart's argument shows that we must abandon our prejudice that there must be a complete [observer-independent] description of reality." (Dummett 1960, p. 504, [1]).

Note that rejection of the 'prejudice' that there must be an observer-independent description of reality sits comfortably with the central idea of Dummett's anti-realist programme in semantics and epistemology (the idea that reality cannot outrun what we can, in principle, know). Finally, we can now see that the title of Dummett's article is something of a misnomer. Dummett is not defending McTaggart's proof of the unreality of time, but arguing that it can be transformed into a (quite different) proof of the falsity of the observer-independence thesis.

Acknowledgements

Thanks to an audience at ANU for useful feedback.

Bibliography

[1] Dummett, M. (1960). A Defence of McTaggart's Proof of the Unreality of Time. *Philosophical Review* **69**(4): 497-504.

[2] McTaggart, J.M.E. (1908). The Unreality of Time. *Mind* **17**(4): 457-474.

[3] Thomson, J.J. (2001). McTaggart on Time. *Noûs* **35**: 299-252.

A Logical Framework for the Spotlight Theory of Time

Ciro De Florio
Catholic University of Milan, Italy
ciro.deflorio@unicatt.it

Aldo Frigerio
Catholic University of Milan, Italy
aldo.frigerio@unicatt.it

Alessandro Giordani
Catholic University of Milan, Italy
alessandro.giordani@unicatt.it

Abstract

The aim of this paper is twofold: on the one hand, to provide a conceptual analysis of the so called moving spotlight theory of time, highlighting its underlying metaphysical assumptions; on the other hand, to develop a semantics for a temporal logic which allows us to grasp the philosophical background of this theory.

Keywords: Temporal logic, Spotlight theory of time, Eternalism, Absolutism, Block universe, Dynamism, C. D. Broad

1 Introduction

The moving spotlight theory of time, MST for short, can be best understood in terms of Broad's lucky metaphor (see [1]):

> We are naturally tempted to regard the history of the world as existing eternally in a certain order of events. Along this, and in a fixed direction, we imagine the characteristic of presentness as moving, somewhat like the spot of light from a policeman's bull's-eye traversing the fronts of the houses in a street. What is illuminated is the present, what has been illuminated is the past, and what has not yet been illuminated is the future. (p. 59, [1])

Ironically, Broad himself did not advocate this view, being a growing block theorist, but the image of the policeman patrolling the neighbourhood became notorious in the successive debate.

It is worth noting that MST is a metaphysically rich theory that assumes both the cumbersome support of all eternal facts and the elusive dynamics of the present. Usually, metaphysical costs are to be justified in the light of high explanatory power and in order to identify this power it is necessary to offer a very precise *conceptual characterization* of the view at stake. In our opinion, in spite of the publication of very good and interesting recent works on this subject [1], such a characterization is still lacking, and so our first aim is to provide both a conceptual analysis of the main tenets of MST and a formal model of the semantics of its temporal operators. Thus, in the first part of this work, we focus on the main metaphysical assumptions of the spotlight theory of time. In particular, we stress its "double" core, which combines elements from a static view of time, namely the assumption of a block-like eternal universe, with elements from a dynamic view of time, namely the assumption of a moving-light present instant. The third section is then devoted to the development of a model of temporal logic rich enough to allow us to account for the dynamics of the present. The logical device we will use to do that is given by what we call dynamical models, i.e. models constituted by a chain of variants of more standard static models. In constructing such models, we remain agnostic about the topology of time, so that our models allow for both forward and backward branching. The main advantage provided by introducing dynamical models is that they enable us to characterize the present and to distinguish two fundamental modalities concerning the past and two modalities concerning

[1] See [2,3,5]

the future. In more details, we introduce a pair of modalities acting as quantifiers on a *thin* past and a *thin* future, and a different pair of modalities acting as quantifiers on a *thick* past and a *thick* future. The basic idea is then roughly the following: to exist in the thin past coincides with being located before the instant of reference, while to exist in the thick past coincides with being located before the instant of reference *and*, in accordance with the spotlight metaphor, to have been illuminated by the light of the spotlight. Furthermore, as far as the future is concerned, things are not entirely symmetrical. In fact, while it is possible for something to exist at a thin future, since to exist at the thin future coincides with being located after the instant of reference, it is problematic to say that something exists at the thick future, since the thick future could be non-existent. In other terms, something *is* actually at the thick future only *retrospectively*, that is, only from the point of view of the past. Indeed, with respect to the actually present instant, there is no thick future, since the present instant is the instant now illuminated by the spotlight, so that no successive instant can have been illuminated.

The present framework has interesting philosophical applications. The first one concerns the possibility of distinguishing two kinds of change. On the one hand, a "static" kind of change, corresponding to the fact that something, an apple say, which is red at a time t_1, is brown at a time t_2 following t_1. On the other hand, a "dynamic" kind of change, corresponding to the fact that something, which is now red, becomes brown, as a consequence of the fact that the present instant itself is changing. A second, natural, application concerns the extension and the adaptation of the dynamical model to other *A*-theories currently debated in philosophy of time, like presentism and the growing block theory. Due to space limitation, we plan to study these applications in further works. Finally, it is important to note that in what follows we are not engaged in a defense of MST. Hence, we will not present any direct argument for the truth of this view of time. However, it is fair to say that many criticisms against MST, and, generally, against similar theories of time, depend on the opacity, or unintelligibility, of such theories. In the light of this, the clarification of the metaphysical basis and the logical structure of MST we are going to provide can be intended as a contribute on the behalf of the advocates of dynamical theories of time.

2 The main features of MST

MST contains elements from the two traditional conceptions of time.[2] From *B*-theory it inherits the structure of the space-time manifold, which is constituted by all present, past, and future facts. From *A*-theory it inherits two elements: the real, objective existence of a tensional trait of the universe, i.e. the present, and its dynamics.

Before starting to investigate these points, a clarification on our construal of the concepts of instant of time, fact, and universe is in order. A block universe can be regarded either as a collection of facts, namely the collection of all existent facts, or as a giant fact, the fact including all existent facts as parts. The difference between these two views depends, in our opinion, on what is considered to be fundamental. Thus, if the universe is assumed to be the fundamental reality, then it seems to be more appropriate to regard individual facts, if there are any, as parts of this fundamental reality. By contrast, if individual facts are assumed to fundamentally exist, then it seems to be more appropriate to regard the universe as the collection of them. We do not want to take a stand on this issue and nothing in what follows depends on that. Once a block universe is given, we can think of time as a certain relational structure connected with it. This connection can be conceived of in two ways. We can assume either that this structure exists independently of the universe, and that different parts of the universe are located at different instants in it, or that the existence of the structure is dependent on the existence of the universe in as much as temporal relations are just relations between parts of the universe. In the first case, time is a substantial entity, with instants of time bearing temporal relations with other instants of time and relations of location with facts, so that two facts are simultaneous precisely when they are located at the same instant. In the second case, time is a relational entity, with parts of the universe bearing temporal relations with other parts of the universe. Still, also in this second case, it is possible to introduce a relation of temporal location by stating that an instant of time is a complete collection of simultane-

[2] A terminological note. We will be rather relaxed about the use of the terms "A-theory" and "B-theory"; generally, we will consider A and B-theories always from a metaphysical point of view – that is, with no reference to the problem of tensed / tenseless propositions – and we will classify the A-theories (or A-views) of time as *dynamical* metaphysical frameworks as opposed to *static* metaphysical frameworks (B-theories).

ous facts and that a fact is located at a certain instant of time precisely when it is part of the collection constituting that instant. Again, we take a stand neither on this issue nor on the metaphysical status of the relation of simultaneity. Accordingly, we will freely use temporal relations as well as relations of location and we will characterize both instants of time and facts as existent and as past, present, and future.

So, we can characterize the spotlight theory as defined according to four basic tenets, respectively concerning: the ontological status of instants of time; the ontological status of the present instant; the substantial aspect of the universe; the dynamical aspect of the universe. In particular:

1. *Eternalism*: all the instants of time, past, present, and future, exist, where the concept of existence is to be intended as univocal and tenseless.[3]

2. *Absolutism*: the instants of time can be divided into past, present, and future in an absolute way; in particular, there is a matter of fact about which instant is present.

3. *Block universe*: the state of the universe is fixed except for what instant of time is the present instant; hence, the identity of the universe as a block is fixed.[4]

4. *Dynamism*: the division of the instants of time into past, present, and future is constantly changing; in particular, what instant of time is the present instant is not fixed.

The richness of MST is not always kindly considered. On the one hand, it is economically disadvantageous both with respect to presentism, since

[3]Here the concept of existence is to be intended in accordance with the main theses proposed in [7]. In sum, there is a unique concept of existence, which (*i*) coincides with the concept of being, (*ii*) is not a concept concerning an activity, and (*iii*) is completely captured by the existential quantifier of first order logic.

[4]This assumption is not shared by all proponents of a spotlight theory of time. See [4] for a version of the theory which, instead of adopting a block universe as a substratum of change, introduces a variant of the growing block universe. It is also worth noting that this assumption is not implied by eternalism, since it is possible for all the instants of time to exist eternally, even though what obtains at such instants changes as time goes by. On this view, there are always past, present and future facts, but they are not always the same facts.

it adds a plethora of non present, but existing, facts, and with respect to the block-universe theory, since it adds the real and dynamical property of presentness. On the other hand, it seems to be theoretically disadvantageous, since, as highlighted by [6], this view "combines the least commonsensical consequence of the B-theory (eternalism) with the most scientifically problematic consequence of the A-theories (the privileged present)". This notwithstanding, in the last few years, two book-length contributions[5] and a series of papers[6] have been produced which are devoted to the analysis and defense of various versions of MST, thus witnessing an increasing interest for this non-standard A-theory of time. However, up to now, no standard version of this theory is available, and the best attempt to characterize our intuitions on the moving spotlight (which is [5] in our opinion) is not sympathetic towards this view. Hence, in the following section, we offer our contribution to this trend by providing a formal model able to grasp the main features of MST.

3 A Model for MST

STATIC AND DYNAMICAL MODELS

In this section, we introduce the temporal language we consider appropriate for describing the dynamics of the present and define the models capturing the structure of a universe with a moving spotlight.

Definition 1. Language for the moving spotlight theory of time.

The language $\mathscr{L}$ of the moving spotlight theory of time contains a countable set $\{p_i\}$ of propositional variables, two *static* temporal operators $\langle P\rangle$ and $\langle F\rangle$, two new *dynamic* temporal operators $\langle P!\rangle$ and $\langle F!\rangle$, and a modality $[now]$. The set $Fm(\mathscr{L})$ of formulas of $\mathscr{L}$ is then defined as follows.

$$\phi ::= p_i \mid \neg\phi \mid \phi \wedge \psi \mid [now]\phi \mid \langle P\rangle\,\phi \mid \langle P!\rangle\,\phi \mid \langle F\rangle\,\phi \mid \langle F!\rangle\,\phi$$

The other propositional connectives and the dual modalities are defined in the standard way. Intuitively, p_i states that a basic state of affairs is true at the time of evaluation, without further assumptions on

[5] See [2] and [5].
[6] See for instance, [3].

the set of basic states of affairs. $[now]\phi$ states that ϕ is true at the actually present instant; $\langle P\rangle\,\phi$ states that ϕ is true at some instant that precedes the instant at which that proposition is evaluated, while $\langle P!\rangle\,\phi$ states that ϕ is true in the *actual past*, that is at some instant that precedes both the instant at which that proposition is evaluated and the *present* instant along the actual history; analogously, $\langle F\rangle\,\phi$ states that ϕ is true at some instant that is preceded by the instant at which that proposition is evaluated, while $\langle F!\rangle\,\phi$ states that ϕ is true in the *actual future*, i.e. at some instant that is preceded by the instant at which that proposition is evaluated but is in the actual part of the history, which is the part of the history travelled by the spotlight.

Definition 2. Temporal Frame.
A temporal frame is a pair $(T,\leq)$, where

(i) $T \neq \varnothing$
(ii) $\leq$ is a reflexive and transitive relation on T.

Intuitively, T is the set of instants, or moments, of time, while $\leq$ is the relation that a instant x bears to a instant y just in case either $x = y$ or x precedes y. It is worth noting that we are not assuming that T is a chain and we allow both forward and backward branching. The models for the language of the spotlight theory of time are based on a temporal frame and are subdivided into *static* and *dynamical* models.

Definition 3. Static model.
A static model is a tuple $M = (T,\leq,\mathscr{P},V)$ where

(i) $(T,\leq)$ is a temporal frame
(ii) $\mathscr{P} = \{t\}$, for some $t \in T$
(iii) $V : \{p_i\} \to \wp(T)$

Intuitively, in a temporal frame, $\mathscr{P}$ is the extension of the property of being present, so that it contains the unique element of T that coincides with the present instant, while V is a function that assigns to propositions the sets of instants at which the propositions hold true.

Definition 4. Present instant in a static model.

The present instant t_M in a static model $M = (T,\leq,\mathscr{P},V)$ is the unique instant of T that is contained in $\mathscr{P}$.

Definition 5. Variant of a static model centered at a time.

A variant of a static model $M = (T, \leq, \mathscr{P}, V)$ is a tuple $M_t = (T, \leq, \mathscr{P}_t, V)$ where $\mathscr{P}_t = \{t\}$, for $t \in T$.

Thus, a variant of M is a model $M_t = (T, \leq, \mathscr{P}_t, V)$ that coincides with M except for the instant that has the property of being present, i.e. except for the *now*. This is in line with both the eternalist and the dynamical traits of the spotlight theory. What changes is the point that is present, the point where the present is located, while all the instants of time are held fixed.

Corollary: $t_{M_t} = t$.

As straightforward consequences of this definition we obtain that (i) every model is a variant of itself, since $M = M_{t_M}$; (ii) if M_{t_1} and M_{t_2} are variants of M, then M_{t_1} is a variant of M_{t_2}, since $M_{t_1} = (M_{t_2})_{t_1}$.

Therefore, being a variant is an equivalence relation and induces a partition on the class of all the static models.

Definition 6. Dynamical model.

A dynamical model $D(\mathfrak{M})$ based on $\mathfrak{M} = (T, \leq, \mathscr{P}, V)$ is a set of variants of $\mathfrak{M}$ satisfying the following conditions:[7]

(i) $\mathfrak{M} \in D(\mathfrak{M})$
(ii) if $M_t \in D(\mathfrak{M})$, then $t \leq t_{\mathfrak{M}}$
(iii) if $M_{t_1}, M_{t_2} \in D(\mathfrak{M})$, then $t_1 \leq t_2$ or $t_2 \leq t_1$

The basic idea here is that a dynamical model based on $\mathfrak{M}$ contains the present model of the universe, which is $\mathfrak{M}$, together with a chain constituted by models of the past of the universe with respect to $\mathfrak{M}$. A dynamic model then represents the course of the universe up to the present time: this course is indeed a chain, but it is represented not as *a chain of instants*, but *as a chain of models*, which are precisely the models of the universe at the instants that have been actually present in the past, thus constituting the actual history up to $t_{\mathfrak{M}}$. In addition, all these models give rise to past dynamical models since, if $M_{t_1} \in D(\mathfrak{M})$, then $D(M_{t_1}) = \{M_{t_2} \in D(\mathfrak{M}) \mid t_2 \leq t_1\}$ is a dynamical model based on M_{t_1}, being constituted by M_{t_1} plus a chain of models for the past of the universe with respect to M_{t_1}.

[7] In what follows $\mathfrak{M}$ is the static model on which the dynamic model is based.

The truth conditions of the propositions of $\mathscr{L}$ are now definable as follows, where M varies over elements of a dynamical model $D(\mathfrak{M})$.

Definition 7. Truth at an instant of M in $D(\mathfrak{M})$.

$M,t \models p_i \Leftrightarrow t \in V(p_i)$
$M,t \models \neg\phi \Leftrightarrow M,t \not\models \phi$
$M,t \models \phi \wedge \psi \Leftrightarrow M,t \models \phi$ and $M,t \models \psi$
$M,t \models [now]\phi \Leftrightarrow M,t_M \models \phi$
$M,t \models \langle P\rangle\,\phi \Leftrightarrow M,t' \models \phi$, for some $t' < t$
$M,t \models \langle F\rangle\,\phi \Leftrightarrow M,t' \models \phi$, for some $t < t'$
$M,t \models \langle P!\rangle\,\phi \Leftrightarrow M_{t'},t_{M_{t'}} \models \phi$, for some $t' < t$
$M,t \models \langle F!\rangle\,\phi \Leftrightarrow M_{t'},t_{M_{t'}} \models \phi$, for some $t < t' \leq t_{\mathfrak{M}}$

The first condition states that p_i is true at t in M just in case p_i is true according to V. Since V is the same for all the models in $D(\mathfrak{M})$, the truth of p_i at t is fixed in all the models in $D(\mathfrak{M})$, in accordance with the intuition that the universe is a fixed block. As to the modal operators:

- $[now]\phi$ is true at t in M precisely when ϕ is true at the present instant of time t_M in M;
- $\langle P\rangle\,\phi$ is true at t in M precisely when ϕ is true at some instant that precedes t;
- $\langle F\rangle\,\phi$ is true at t in M precisely when ϕ is true at some instant that is preceded by t;
- $\langle P!\rangle\,\phi$ is true at t in M precisely when ϕ is presently true in some variant of $\mathfrak{M}$ that precede M_t in the dynamical model;
- $\langle F!\rangle\,\phi$ is true at t in M precisely when ϕ is presently true in some variant of $\mathfrak{M}$ that lies between M_t and $\mathfrak{M}$ in the dynamical model.

Importantly, the truth of $\langle P!\rangle\,\phi$ implies that $[now]\phi$ has been true in some instant of the actual past of M, and hence of $\mathfrak{M}$, and the truth of $\langle F!\rangle\,\phi$ implies that $[now]\phi$ has been true at some instant of the actual future of t in M, and hence again in the actual past of $\mathfrak{M}$, since the *actual* instants lie in the past of the present instant of $\mathfrak{M}$.

Note 1 The significance of the distinction between standard temporal modalities and bang temporal modalities becomes evident if we consider the truth-conditions of $\langle P \rangle \phi$ and $\langle P! \rangle \phi$. While $\langle P \rangle \phi$ is true at t in M just in case there is a *possible* past instant, relative to t, at which ϕ is true, $\langle P! \rangle \phi$ is true at t in M just in case there is an *actual* past instant, relative to t, at which ϕ is true. Therefore, in terms of the spotlight metaphor, while the truth of $\langle P \rangle \phi$ is determined by the state of the block universe, the truth of $\langle P! \rangle \phi$ is determined by the *path* followed by the light through the universe.

Note 2 The present framework provides us with the resources to express the possible indeterminateness of the past and the future with respect to a dynamic model $D(\mathfrak{M})$.

Definition 8. Being indeterminate in M at t.

(i) in M, the past is indeterminate at $t := M, t \models \neg \langle P! \rangle \top$
(ii) in M, the future is indeterminate at $t := M, t \models \neg \langle F! \rangle \top$

Thus, it is always true that the future is indeterminate at the present instant, since $\mathfrak{M}, t_{\mathfrak{M}} \models \neg \langle F! \rangle \top$, given that there exists no variant of $\mathfrak{M}$ in $D(\mathfrak{M})$ which is indexed at an instant that is in the future of $t_{\mathfrak{M}}$.

Note 3 In accordance with our intuition on the structure of the dynamical universe, there exists no actual future for the present instant. This condition allows us to introduce the following classification of instants.

(1) in M, t is actually past $:= M, t \models \langle F! \rangle \top$
(t is actually past: it has an actual future)

A time in the actual past is a time with respect to which there is an actual future, since the future is not determinate for both the present and every time that follows the present.

(2) in M, t is in the future $:= M, t \models \langle P \rangle \neg \langle F! \rangle \top$
(t is in the future: it comes after the actual present)

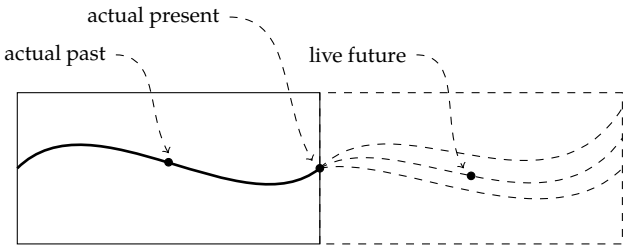
A time in the future is a time that follows the present. Still, for every time t that follows the present there is a past instant, between the present and t, which has no determinate future.

(3) in M, t is actually present $:= M, t \models \neg \langle F! \rangle \top \wedge \neg \langle P! \rangle \neg \langle F! \rangle \top$

(t is actually present: it has an actual past, but no actual future)

The present time $t_{\mathfrak{M}}$, like any time that is not in the actual past, has no determinate future. In addition, the present time is the unique time that is preceded by all the times with respect to which there is an actual future. As expected then, the present turns out to be definable as the unique instant that lies on the edge of the actual history, so that the previous proposition uniquely identify $t_{\mathfrak{M}}$ in the dynamical model.

The following picture synthesizes these definitions.



Note 4 A last significant upshot is the possibility of expressing the fact that something is a lost possibility.

Definition 9. Being a lost possibility, in M at t.

We define ϕ to be a lost possibility, in M, at t, if and only if:

$$M,t \models \langle P! \rangle \langle F \rangle \phi \wedge \neg \langle P! \rangle \phi \wedge [now] \neg \langle F \rangle \phi$$

Hence, in the present framework, a formula like $\langle P \rangle \langle F \rangle \phi \wedge \neg \langle F \rangle \phi$ is not appropriate for expressing that ϕ is a lost possibility, since an instant of time can be in the future of a past instant without being in the future of an instant which is in the actual past of the instant of evaluation. By contrast, a formula like $\langle P! \rangle \langle F \rangle \phi \wedge \neg \langle P! \rangle \phi \wedge [now] \neg \langle F \rangle \phi$ states that, in the actual past, ϕ could have obtained, even if, at the instant of evaluation, it has not obtained and it is no longer the case that it can obtain.

4 Conclusion

The main contributions of this paper can be summed up as follows.

1. To put forward a logical characterization of MST by providing a formal model of it.

2. To put forward a precise characterization of the concept of presentness and its dynamics.

3. To show how some basic notions in the philosophy of time can be defined in this framework.

In a future work, the main results of this paper could be extended to other A-theory of time such presentism or growing block theory. Of course, appropriate changes are needed in order to deal with the different ontology of facts of these A-theories.

Acknowledgements

The authors thank Patrick Blackburn, David Jakobsen, Peter Øhstrøm, and an anonymous referee for very useful comments on earlier versions of this paper.

Bibliography

[1] Broad, C.D. (1923). *Scientific Thought*. London: Routledge & Kegan Paul.

[2] Cameron, R.P. (2015). *The Moving Spotlight*. Oxford: Oxford University Press

[3] Deasy, D. (2015). The Moving Spotlight Theory. *Philosophical Studies* **172**: 2073–2089.

[4] Skow, B. (2012). Why does Time Passes?. *Noûs* **46**: 223–242.

[5] Skow, B. (2015). *Objective Becoming*. Oxford: Oxford University Press.

[6] Sullivan, M. (2016). Review of Ross P. Cameron: The Moving Spotlight: An Essay on Time and Ontology. *Notre Dame Philosophical Review* **2016.08.26**. https://ndpr.nd.edu/news/the-moving-spotlight-an-essay-on-time-and-ontology/. Accessed Dec 12, 2018.

[7] van Inwagen, P. (2009), Being, Existence, and Ontological Commitment. In: Chalmers, J., Manley, D. and Wasserman, R. (eds.): *Metametaphysics. New Essays on the Foundations of Ontology*. Oxford: Oxford University Press: pp. 472–506.

Now There Will be Trouble

Giuseppe Spolaore
Università di Padova, Italy
giuseppe.spolaore@gmail.com

Fabio Del Prete
CNRS, France
fab.delprete@gmail.com

Abstract

We consider sentences in which "now" occurs in initial position and show that the meaning they convey differs from the meaning of sentences that are otherwise identical except for "now" occurring in final position. We argue that the occurrence of "now" in initial position triggers a particular kind of modal reading for the sentence to which the adverb is prefixed. A general notion of modal forcing is proposed to provide a uniform account of this kind of reading. Armed with this account, we offer a solution to two tense-modal puzzles, which have to do with fatalism and the possibility of a changing past.

Keywords: Now, Modal forcing, Time adverbs, Tense and modality, Tense-modal puzzles, Fatalistic argument, Changing past.

1 Two little puzzles

It is September 13, 2017. Paris has just been declared host city of the 2024 Olympic Games by the International Olympic Committee. Right after the declaration, you can felicitously and truthfully utter (1):

(1) Now Paris will host the 2024 Olympic Games.

Let t be the time of your utterance. By uttering (1) at t, it seems that you convey the implicature that (2) was false at some time preceding t:[1]

(2) Paris will host the 2024 Olympic Games.

Let us observe, however, that (1) entails (2). As a consequence, (2) must also be true at t. From this, assuming the platitudes (P_1) and (P_2) about the meaning of (2) and the natural ordering of times, we can conclude that (2) was true at any time before t.

(P_1) Sentence (2) (considered at time t) says that Paris hosts the 2024 Olympic Games at some time in the future (relative to t).

(P_2) If a time is in the future of t, then it is in the future of any time preceding t.

Therefore, your utterance of (1) at t both implicates that (2) was sometimes false before t *and* entails that (2) was always true before t. In other terms, what your utterance implicates clashes with what your utterance entails.

Notice that this clash should make your utterance of (1) infelicitous, for essentially the same reason that (3) sounds infelicitous:

(3) ??Mary has two children. She has given birth to one boy and two girls.

Arguably, what is wrong with the discourse (3) is that one of its implicatures (the scalar implicature that Mary has *exactly* two children) is at odds with something that (3) entails (that Mary has at least three children). In contrast, your utterance of (1) is perfectly felicitous. How is this possible?

A similar puzzle arises with a sentence about the past. Consider the following scenario. An *ex post facto* law $\mathscr{L}$ is enforced at t; $\mathscr{L}$ classifies certain actions as criminal, while the same kind of actions were previously regarded as lawful. Moreover, Bill had performed such an action on a certain day d, before t. At t, as law $\mathscr{L}$ is enforced, you can felicitously and truthfully utter (4):

[1] When we speak of a sentence as being true or false at a time, we only mean to say that the sentence is true or false *as evaluated* at that time. All we say is meant to be consistent with a tenseless conception of (propositional) truth.

(4) Now Bill committed a crime on day d.

Your utterance of (4) entails (5):

(5) Bill committed a crime on day d.

Therefore, (5) must also be true at t. From this, assuming the platitudes (P_3) and (P_4) about the meaning of (5) and the natural ordering of times, we can conclude that (5) was true at any time between day d and t.

(P_3) Sentence (5) (considered at time t) says that Bill commits a crime at some time in the past (relative to t), which falls within day d.

(P_4) If a time t' is in the past of t, then it is in the past of any time between t' and t.

Moreover, your utterance of (4) at t conveys the implicature that (5) was false at some time between day d and t. Therefore, your utterance of (4) has an entailment and an implicature that contradict each other. This should make your utterance infelicitous, contrary to the facts. Again, how is this possible?

We believe that the puzzle about the future has a very natural solution – so natural that it was often anticipated by our audiences at previous presentations of this work. The solution is based on the following observation: sentences about the future sometimes have a reading involving some planning (a *plan reading*, for short) , that is, they are used to talk about events that are expected to occur in the future given certain present plans. Prototypical examples of such sentences are the so-called "futurates" [see, e.g., 9, 8], such as:

(6) Paris hosts the 2024 Olympic Games.

(7) Paris is hosting the 2024 Olympic Games.

However, the existence in English of dedicated tense-aspect forms to express a plan reading, as in (6)–(7), does not exclude that future tense sentences like (2) can occasionally convey the same kind of reading. Now, if (2) does convey a plan reading in the context of (1), what (2) says in that context can be expressed as in (2′):

(2′) There is an official plan according to which Paris is to host the 2024 Olympic Games.

Clearly, (2′) can be false at t even if there is a time following t at which Paris hosts the 2024 Olympic Games.[2] What your utterance of (1) implicates, in the plan reading of its component sentence (2), is that (2′) (and not the proposition that (2) expresses according to (P_1) was false at some time before t. In other words, your utterance of (1) at t implicates that at some time t_0 before t it was false that there was a (then valid) official plan according to which Paris was to host the 2024 Olympic Games. Importantly, this implicature is perfectly consistent with (2) being true at t_0 in the reading given in (P_1), for if (2) was true at t_0 in *this* reading, then at t_0 it was true that Paris would host the 2024 Olympic Games – although this may have been unknown to everybody at the time.

Now, the puzzle about the past cannot be solved in exactly the same way as the puzzle about the future, since it makes no sense to invoke plan readings for sentences about the past. Still, we believe that the two puzzles have the same *kind* of solution. The idea is as follows. What (5) says in the context of (4) is not what (P_3) states that it says; instead, what (5) says in that context can be expressed as in (5′):

(5′) There is an institutional frame according to which Bill committed a crime on day d.

Assuming this reading of (5) (call it *institutional reading*, for short), what your utterance of (4) at t implicates, in the institutional reading of its component sentence (5), is that (5′) (and not the proposition that (5) expresses according to (P_3)) was false at some time between day d and t. In other words, your utterance of (4) at t implicates that at some time t_0 before t it was false that there was a (then valid) institutional frame according to which Bill had committed a crime on day d. Notice that this implicature is perfectly consistent with (5) being true at t_0 in the reading given in (P_3): at t_0 it was true that Bill had committed a crime on day d.

2 Forcing readings and "now"-initial sentences

The general idea we pursue in this section is that plan readings and institutional readings are just a special case of a more general kind, which

[2] It can be added that (2′) can be true at t even if there is no time following t at which Paris hosts the 2024 Olympic Games. See, e.g., [4].

we call *forcing readings*. The fundamental notion is that of a *forcing relation* holding between a state *s* and an event *e*: a relation of determination, whereby *e* occurring at some time is necessitated by *s* holding at another time. Next we consider some linguistic data with the aim of showing that sentences in which "now" contributes to the expression of a forcing reading – like (1) and (4) from section 1 – share a particular syntactic property, formally signaling that "now" is playing a different role here from that of a regular temporal adverb (i.e., one which locates an eventuality in time).

LINGUISTIC PROPERTIES OF "NOW"-INITIAL SENTENCES

If one looks at dialogic contexts in which the addressee denies what the speaker has just said, it is clear that the syntactic position of "now" can make an important difference to the interpretation of an utterance, and consequently to the coherence of a dialogue. An example of this is the contrast between (8) and (9):

(8) [Context: A and B are watching the 2011 World Championships in Athletics. A believes that Yohan Blake will run in a moment.]

A. Yohan Blake will win (right) now.

B. No, that's false. He'll run tomorrow afternoon. / ??Walter Dix could make it as well.

(9) [Context: A and B are watching the 2011 World Championships in Athletics. They start talking just after Usain Bolt has been disqualified for a false start.]

A. Now Yohan Blake will win.

B. No, that's false. Walter Dix could make it as well. / ??He'll run tomorrow afternoon.

In dialogue (9), unlike in (8), the mere *possibility* that Yohan Blake does not win is sufficient for B to deny A's statement. That would be unexpected unless A's statement was understood as having a necessity modal force (viz., in accordance with the forcing reading of the prejacent of "now"). The contrast between (8) and (9) shows that the syntactic position of "now" in the sentence matters for the interpretation: when it contributes to expressing a forcing reading, "now" typically

occurs in sentence-initial position.[3] In what follows, we shall refer to sentences such as (1), (4) and A's statement in dialogue (9) as *"now"-initial sentences*.[4]

It has been observed by many that "now" has a contrastive value (at least when used with predicates that can be true of extended intervals – e.g., with stative predicates; see [10, 2, 14], among others). Mostly, this has been observed for occurrences in which "now" functions as a regular temporal adverb modifying a verb phrase. For instance, an utterance of the sentence "I am tired now" implicates that the speaker was not tired before her utterance.[5] In line with this widespread observation, we notice that "now"-initial sentences have a contrastive flavour: sentence (10) implicates that, at a previous time, no wedding plan existed, which forced the event of John and Mary getting married in April.

(10) [Said in January, just after the couple's wedding plan has changed.]
Now John and Mary are getting married in April.

Analogously, (11) implicates that a previous literary setting forced the event that Holmes died in the Reichenbach falls:

(11) [Said after Holmes 'resuscitated' in *The Adventure of the Empty House*.]
Now Sherlock Holmes didn't die in the Reichenbach falls.

Analogously, again, A's statement in dialogue (9) above implicates that,

[3]It can also occur in sentence-final position, although in this position one needs the right intonation in order to make it clear that the adverb is not being used as a modifier of the verb phrase (hence, as a regular time adverb).

[4]The semantic relevance of the position of time adverbs in the linear order of the sentence, in particular the emergence of modal meanings with pre-verbal occurrences, is a well-known phenomenon, one which has been described and amply documented in the linguistic literature ([20], among many others).

[5]Such implications presumably arise via pragmatic inference – possibly exploiting Grice's maxim of quantity – and are not limited to "now": it seems safe to say that the use of any time adverb to specify the temporal location at which a certain eventuality E holds triggers the inference that E does not hold at other locations, e.g., the sentence "The shop is closed today", without further indications, suggests that the shop was not closed yesterday and will not be closed tomorrow. Said this, we think that there likely is something special to the contrastive implications that "now" gives rise to in "now"-initial sentences, as we discuss in greater detail in section 2.

at a time preceding Usain Bolt's disqualification, there was no state that causally forced the event of Yohan Blake winning.

VARIETIES OF FORCING

The discussion so far should have made it clear that forcing relations come in many varieties. The kind of forcing involved in (1), and in (10) as well, we shall call *forcing from plans*. A relation of forcing from plans obtains between the state of a certain plan being effective and the events that must obtain if the plan is realized. The kind of forcing involved in (4), we shall call *institutional forcing*. Yet another kind of forcing, at play in the interpretation of A's statement in dialogue (9) above, is *causal-historical forcing*: this obtains between a state s and an event e if s causally necessitates e, given certain historical conditions [17]. In both causal-historical forcing and forcing from plans, the temporal ordering between s and e is such that s precedes e (causes precede effects, plans precede the events that they intend). But it is also possible that the temporal relation between e and s is reversed, as it may happen with institutional forcing.[6]

ANALYSIS OF "NOW"-INITIAL SENTENCES

We propose that "now" in the "now"-initial sentences considered above is best modelled as a modal operator N with the following semantic and pragmatic properties:

Stative Anchor: N is anchored to a state s^* which obtains at reference time t_R. State s^* can be described by a *that*-clause argument of "now"[7] and is presupposed (i.e., it is common knowledge that s^* obtains at t_R).

Forcing: N triggers a forcing reading for its prejacent, whereby a contextually relevant forcing relation $\mathscr{R}$ is required to hold between

[6] The kind of forcing involved in (11) seems to be different from all the others considered in the main text: in this case, it is not clear whether s and e can be said to be related via a temporal relation.

[7] The possibility to describe the underlying state via a *that*-clause attached to "now" is exemplified by the more verbose variant of (1) given in (i):

(i) Now *that the International Olympic Committee has made its plan*, Paris will host the 2024 Olympic Games.

See [6].

s and the event e described by the prejacent. (A relation $\mathscr{R}$ holding between a state s and an event e in a world w at a time t is a *forcing relation* when e occurs in every alternative to w at t which is compatible with s.)

Change of State: The evaluation of an utterance of $\ulcorner$now (that S_1), $S_2\urcorner$ at t_R generates an implicature concerning the behaviour of the forcing relation $\mathscr{R}$ at alternative times t' for t_R, to the effect that $\mathscr{R}$ does not hold between any state obtaining at t' and the event e.

This analysis predicts that (1) is true in the context described at the beginning of the paper at the condition that an event of Paris hosting the Olympic Games occurs in the future of September 13, 2017 in every world compatible with a presupposed state s^*, where s^* is a state in which the International Olympic Committee has made its relevant plan. In this case, the variable $\mathscr{R}$ refers to the particular kind of forcing exerted by plans. A completely parallel calculation can be provided for the truth conditions of (4) (understood as "now *that the relevant law is in force* Bill committed a crime on day d"), but in this case $\mathscr{R}$ will refer to institutional forcing.

Let us comment on the Stative Anchor and Change of State properties formulated above. Concerning Stative Anchor, we note that the state s^* which is mentioned in it is typically established in surrounding discourse, as in the following example (from a newspaper article):

(12) [Paris] deputy mayor Bruno Julliard announced that glass panels will replace the grills that are currently weighed down by hundreds of thousands of padlocks. [...] As in many other cities, couples lock their padlocks to bridges and monuments, symbolizing their union. Recently, cities like Melbourne and New York have removed locks from their bridges, and *now Paris will join them*.

The sentence in italics in (12) is understood as "now that glass panels will replace the grills in question (as announced by the deputy mayor), Paris will join those other cities," where the material in the *that*-clause is clearly related to a previous stretch of discourse.

Turning to Change of State, we note that this property mentions *alternative times* for t_R, which are moreover claimed to be *in the past* of t_R. Thinking about the contrastive value of "now" (which was mentioned

above) more broadly – for instance, thinking about the implications of contrast of a sentence like "I am hungry now" – one may wonder whether the alternative times for t_R must necessarily be *in the past* or can in some cases be *in the future* of t_R. We think that the specific operator N found in "now"-initial sentences is lexically specified in such a way that the alternatives for t_R are times *preceding* t_R. Notice that a different operator N^* exists in English, which is built with "now" but is more complex than the plain "now" considered so far, and N^* is such that the alternatives for t_R relevant for the interpretation of sentences containing N^* are times *following* t_R. The operator in question is syntactically realized as "for now". An example of it is given in (13), which sharply contrasts with (14) in its implications:

(13) For now I have no job.

(14) Now I have no job.

On the one hand, (14) implies that the speaker has become unemployed, that is, a state holds at the present time t which makes it the case that the speaker is unemployed at t but no state held at any time preceding t which made it the case that the speaker was unemployed then. On the other hand, (13) implies that the speaker may end up getting a job. We contend that both prefixes "for now" and "now" in (13) and (14) above use alternative times (for the reference time t_R), and they both require that some relation does not hold at those alternative times. The difference between those two prefixes is that "for now" *looks forward in time* to find its alternatives, while "now" *looks backward*; from these opposite temporal orientations of the two operators, the different implications of (13) and (14) follow: simplifying somewhat, (13) implies that the speaker does not have a job at present but could be no longer jobless in the future, while (14) implies that the speaker does not have a job at present but was not jobless in the past. In what follows we will not elaborate on this difference of temporal orientation between the modal operator "now" (of "now"-initial sentences) and linguistically related operators.[8]

[8]These remarks highlight the importance of taking temporal orientation into account in order to develop a semantic theory of "now" and related modal operators; moreover, they help framing the discussion of the relevant uses of "now" within the broader context of temporal orientation of modals in natural language [7].

In the next section we show that our analysis allows for a treatment of two philosophical puzzles involving time and necessity: the fatalistic argument, first discussed by Aristotle, and a more recent argument concluding to the possibility of changing the past, presented in [3].

3 Two puzzles

THE FATALISTIC PUZZLE

There are a few philosophical arguments that allow one, starting from *prima facie* plausible premises, to draw the fatalist conclusion that the future is settled, that is, historically necessary. These arguments have puzzled generations of philosophers over the centuries and still today keep their grip on the minds of those who seek to defuse them. Puzzlement stems from two sources: first, fatalism in itself is a nearly incredible doctrine; second, it is surprising to see that fatalism can be justified on purely logico-linguistic grounds. Here we shall focus on a very simple argument, which Aristotle discusses in his *De Interpretatione* (19a23-25) (see [21] for an alternative reconstruction of the argument).

Consider the following sentence, which intuitively concerns a contingent eventuality:

(15) There will be a sea battle tomorrow.

By the principle of bivalence, (15) is either true or false. Let us start by assuming that (15) is true. If so, then (15) is true *now*, that is:

(16) Now it is true that there will be a sea battle tomorrow.

But if it is already true now that there will be a sea battle tomorrow, then the battle is fated, that is, inevitable – it is always too late to change the present. Moreover, what is inevitable is not contingent. Therefore,

(17) It is not contingent that there will be a sea battle tomorrow.

We run into a similar conclusion if we assume that (15) is false. Since this argument does not rely on any specific feature of the example chosen, it can be generalized. The conclusion is that everything is either inevitable or impossible: fatalism is true.

Many philosophers accept the step from (15) to (16) as valid and reject the step from (16) to (17) as invalid – this is [12]'s solution to the fatalistic argument [see also 5, 13, 11]. Our analysis of *now*-initial sentences strongly suggests the opposite stance: the step from (16) to (17) is valid, and the invalid step is the one from (15) to (16) [see [16] for a similar diagnosis]. It is natural to think that, if a sentence of the form ⌜Now P⌝ receives a forcing reading in a certain context, then also the corresponding sentence ⌜Now it is true that P⌝ receives a forcing reading in that context.[9] But if (15) does not validly entail (16), then, for the very same reason, (15) does not validly entail the following:

(18) Now there will be a sea battle tomorrow.

Indeed, as uttered in the relevant context, (18) entails that a causal forcing relation exists between some present state and the future event of a sea battle tomorrow. Nothing similar holds for (15).

Before considering our second puzzle, let us note that a perfectly respectable version of the fatalistic argument can be obtained by replacing (18) with a sentence in which "now" does not occur in initial position, for instance:

(19) It is true now that there will be a sea battle tomorrow.

Moreover, and perhaps more importantly, it is easy to envisage a variant of the fatalistic argument in which the role of (18) is taken by a past-tensed truth ascription, for instance:

(20) Yesterday it was already true that there would be a sea battle tomorrow.

If our solution to the fatalistic puzzle is correct, the existence of these alternative arguments suggests that forcing readings can be triggered by a wide array of linguistic constructions, which include but are not limited to "now"-initial sentences. In turn, if this conclusion is correct, then our approach to forcing is in wait of substantial generalization. We shall briefly discuss this prospective generalization in section (4).

[9] Strictly speaking, it is the prejacent of "now" which receives a forcing reading. We can also speak – in a derivative sense – of 'forcing reading' of a "now"-initial sentence, in those cases in which the prejacent receives a forcing reading.

THE PUZZLE OF THE CHANGING PAST

The second puzzle that we wish to consider is related to an argument proposed by [3], calling into question the view that the past cannot change. One way to present this argument is by making use of sentence (21),[10] uttered in the real-world context described below:

(21) [Context: On July 23, 2000, being the rider with the lowest overall time at the end of the last stage, Lance Armstrong had been declared the winner of the Tour de France by Union du Cyclisme Internationale (UCI). It is now October 22, 2012: having discovered that Armstrong made use of banned substances, UCI withdraws all of Armstrong's wins at the Tour de France.]
Armstrong has no longer won the 2000 Tour de France.

Barlassina and Del Prete remark that (21) is a true sentence about the past, moreover, for (21) to be a true sentence about the past, it must be the case that both (i) and (ii) hold true:

(i) at some t before now, the past was such that Armstrong won the 2000 Tour de France;

(ii) now the past is such that Armstrong did not win the 2000 Tour de France.

The conjunction of (i) and (ii) entails that the past has changed in the passage from t to now and so sounds puzzling, since common sense is strongly at odds with the view that the past can ever change.

If our analysis of "now"-initial sentences is correct, we have a principled way out of this puzzle. First of all, observe that (21) essentially conveys the same content as:

(22) Now Armstrong has not won the 2000 Tour de France.

Both (21) and (22) naturally receive a forcing reading. In other words, both say that a presently obtaining state (i.e., the state containing the presently effective declarations by UCI) institutionally forces the eventuality that Armstrong did not win the 2000 Tour de France. Moreover,

[10] Barlassina and Del Prete's main argument is actually based on an intuition about the change of truth value of the context-insensitive sentence about the past "Lance Armstrong won the Tour de France in 2000" across two subsequent real-world contexts.

both (21) and (22) imply that things were different in the past as regards Armstrong not having won the 2000 Tour de France. But if our analysis is on the right track, this implication can be dealt with without assuming that the past has actually changed: we only have to recognize (a) that an institutional event (such as an event of winning the Tour de France) depends on the assignment of a status function by a competent authority [16, 18, 1], and (b) that a state obtained at some time in the past, which institutionally forced the event of Armstrong winning the 2000 Tour de France.

4 Perspectives for future work

In this paper, we presented a specific linguistic phenomenon, the forcing reading of "now"-initial sentences, for which we briefly and informally described a possible treatment. In our proposal, a "now"-initial sentence says that some presupposed state s^*, obtaining at reference time t_R (usually, the present), forces a certain event e, the relation of forcing $\mathscr{R}$ being one of a number of possible relations of determination. Moreover, such a sentence implicates that, for some alternative time t' (typically before t_R), $\mathscr{R}$ does not hold between any state obtaining at t' and the event e. In the last part of the paper, we put our treatment at work on two philosophical puzzles. We are aware that a lot of things remain to be done in this area. In this final section we hint at some perspectives for future work. In passing, we also address some natural perplexities that this too-brief presentation might raise.

Firstly, of course, we have not provided any formal definition of our proposal. We are pretty sure, however, that this can be done, and we plan to offer a formal semantic-pragmatic analysis of forcing sentences in the future (stay tuned!).

Secondly, we mentioned that forcing readings can also be conveyed by sentences involving the truth predicate, such as "It is true now that there will be a sea battle tomorrow". It is interesting to observe that, by using the truth predicate or related constructions, we can produce forcing statements about *past* or *future* presupposed states. For instance, consider (18) again ("Now there will be a sea battle tomorrow") and assume it was uttered yesterday at 3pm. Now suppose that today, in the midst of the battle, you want to reassert the same content – including the implied condition that the prejacent of "now" was false at some

previous time. You could not say:

(23) *Yesterday at 3pm there would be a sea battle today.

Arguably, the best you could do is to say something along the following lines:

(24) a. Yesterday at 3pm it was already true that there would be a sea battle today.

b. Yesterday before 3pm it became true that there would be a sea battle today.

c. Yesterday at 3pm it was already the case that there would be a sea battle today.

A similar strategy applies to all "now"-initial sentences. Therefore, it should be possible to generalize the proposal carried out here (or, better, a suitable formalization thereof) to forcing sentences involving past or future presupposed states. Again, this generalization is left for future work.

We hope to have shown both the philosophical and the linguistic interest of the phenomenon that we have highlighted and of the proposed treatment. The phenomenon supports the view that modality in natural language is a pervasive feature, affecting the interpretation in context of utterances of linguistic structures of which, *a priori*, one would not think they are modal. In particular, no one of our "now"-initial sentences contain overt modal verbs (with the only possible exception of (1), if the future tense auxiliary "will" is to be regarded as a modal verb). The phenomenon also confirms the importance of syntactic structure for the emergence of modal meanings: we have shown that the occurrence of "now" in initial position is a structural pre-condition for "now" to be promoted from the status of regular time adverb (syntactically, a VP modifier; semantically locating an event in time) to the status of a complex presuppositional modal operator. Our data and analysis connect with recent and ongoing researches on dyadic uses of "now" and the importance of presupposed underlying states in the interpretation of natural language [6, 15].

Acknowledgements

We thank Patrick Blackburn, Alessandro Zucchi, Michel Aurnague, Luca Barlassina, Andrée Borillo, Jesse Tseng, as well as the participants to the 4^{th} IAPT conference (Gargnano sul Garda, June 12-14 2017) and to the workshop "Logic and Philosophy of Time: Themes from Prior" (Copenhagen, November 22-24 2017) for their helpful comments.

Funding

This research has been partly funded by a grant from the University of Padova, fund CPDA145940 (2014), based on the project *Tracing the History of Words. A Portrait of a Discipline Through Analyses of Keyword Counts in Large Corpora of Scientific Literature* (P.I. Arjuna Tuzzi, University of Padova).

Bibliography

[1] Khalidi, A.K. (2013). Three kinds of social kinds. *Philosophy and Phenomenological Research* **90**: 96–112.

[2] Altshuler, D. (2016). *Events, States and Times. An essay on narrative discourse in English.* Warsaw/Berlin: De Gruyter Open.

[3] Barlassina, L. and Del Prete, F. (2015). The puzzle of the changing past. *Analysis* **75**(1): 59–67.

[4] Bonomi, A. and Del Prete, F. (2007). Evaluating future-tensed sentences in changing contexts. Manuscript.

[5] Braüner, T., Hasle, P. and Øhstrøm, P. (2000). Determinism and the origins of temporal logic. In: Gabbay, D., Fisher, M.H. and Barringer, H. (eds.): *Advances in temporal logic*. Springer: pp. 185–206.

[6] Carter, S. and Altshuler, D. (2017). 'now' as a dyadic connective. Abstract for the 27th *Semantics and Linguistic Theory* Conference.

[7] Condoravdi, C. (2002). Temporal interpretation of modals. In: Beaver, D., Kaufmann, S., Clark, B. and Casillas, L.(eds.): *The Construction of Meaning*. CSLI Publications: pp. 59–88.

[8] Copley, B. (2009). *The Semantics of the Future*. New York: Routledge.

[9] Dowty, D. (1979). *Word meaning and Montague grammar: the semantics of verbs and times in generative semantics and in Montague's PTQ*. Dordrecht: D. Reidel Publishing Company.

[10] Hunter, J. (2010). *Presuppositional indexicals.* Ph.D. Dissertation, University of Texas, Austin.

[11] Malpass, A. and Wawer, J. (2012). A future for the thin red line. *Synthese* **188** (1): 117–142.

[12] Ockham, W. (1978). *Tractatus de praedestinatione et de praescientia Dei respectu futurorum contingentium*. N.Y.: St. Bonaventure University. Trans. in: Adams, M. and Kretzmann, N. (eds.): *William of Ockham: Predestination, God's foreknowledge, and future contingents*. 2nd edition. 1981. Hackett.

[13] Øhrstrøm, P. (2009). In defence of the thin red line: A case for Ockhamism. *Humana.Mente* **8**:17–32.

[14] Recanati, F. (2004). *Indexicality and context shift.* Unpublished manuscript, Harvard University.

[15] Schwarzschild, R. (2017). Underlying states. Handout of a class given at ESSLLI 2017.

[16] Searle, J. (1995). *The Construction of Social Reality*. New York: Free Press.

[17] Thomason, R.H. (1984). Combinations of tense and modality. In: Gabbay, D. and Guenthner, F. (eds.): *The Handbook of Philosophical Logic, Vol. 2*. Dordrecht: Reidel: pp. 135–165.

[18] Thomasson, A. (2003). Foundations for a social ontology. *ProtoSociology* **18-19**:269–290.

[19] Tooley, M. (1997). *Time, Tense, and Causation*. Oxford: Oxford University Press.

[20] Traugott, E.C. and Dasher, R. B. (2001). *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

[21] Whitaker, C.W.A. (2002) *Aristotle's* De Interpretatione*: Contradiction and Dialectic*. Oxford: Oxford University Press.

Actuality and Possibility in Branching Time: The Roots of Transition Semantics

*Antje Rumberg**
Stockholm University, Sweden
antje.rumberg@philosophy.su.se

Abstract

The transition semantics presented in Rumberg (2016, [5]) constitutes a novel approach to Prior's theory of branching time. It has its roots in the peculiar interrelation of actuality and possibility that we find in branching time structures: the forking paths in the tree of possibilities represent temporal alternatives for a dynamic actuality to evolve rather than modal alternatives to a given actuality. In this paper, we discuss the peculiar interrelation of actuality and possibility that is at the heart of the theory of branching time and illustrate how this interrelation is reflected in the transition semantics.

Keywords: Branching Time, Transition Semantics, Open Future, Actuality, Possibility, Contingency, Stability.

1 Introduction

Possibilities are often conceived of as representing modal alternatives to actuality, i.e. ways our world could be in a certain respect but actually

*This paper is partly based on my PhD thesis (Rumberg 2016, [6]).

is not. We may, for instance, say that it is possible that I am in Copenhagen right now, while I am actually sitting in Konstanz at my desk typing these sentences. The idea of modal alternatives famously underlies the possible worlds approach. Each possible world stands in for some possibility, and one of them is the actual world, the world we actually live in.

The theory of branching time, pioneered by Prior (1967, [4]), affords a different understanding of possibilities. Branching time structures depict the future as genuinely open. The picture is that of a tree whose branches represent possibilities for the future. At the moment, I am sitting here in Konstanz at my desk typing these sentences, and it is possible that I will continue writing for another few hours. But there are alternate possibilities. I can take a break from writing and go for a swim in the Lake Konstanz. And I may just as well head to the airport and take the next flight to Copenhagen. These scenarios all constitute genuine possibilities for the future: as of now, any of them can be actualized.

The relation between actuality and possibility that is at play here is a rather peculiar one: possibilities represent temporal alternatives for a dynamic actuality to evolve rather than modal alternatives to a given actuality. When we are facing the future, we are facing a maze of forking paths, and actuality makes its way through that maze. At each junction, actuality traces one of the possible paths leading toward the future, and all paths are equally viable. At each branching point, one possibility is actualized, and the remainder fades away.

The transition semantics presented in Rumberg (2016, [5]) constitutes a novel approach to Prior's theory of branching time. It has its roots in the peculiar interrelation of actuality and possibility that we find in branching time structures. The distinctive feature of the transition semantics is that it builds on local future possibilities, viz. transitions, rather than on complete possible courses of events. In doing so, it enables a dynamic representation of the interrelation of actuality and possibility and brings to the fore the idea that contingencies about the future dissolve as time progresses.

The paper is structured as follows: in section 2, we introduce the theory of branching time. In section 3, we discuss the peculiar interrelation of actuality and possibility that we find in branching time structures. In section 4, we illustrate how this interrelation bears on the seman-

tics of sentences about the future and motivate the transition semantics. Section 5 provides a formal precisification of the fine-grained picture underlying the transition approach. In section 6, we show how the contingency of the future is dealt with in the transition semantics.

2 A World of Possibilities

Welcome to a world of possibilities! Prior's theory of branching time invites you to a world that is full of possibilities: possibilities for what the future may bring. The future is open. In this section, we briefly introduce the theory of branching time and outline two alternative ways of representing possibilities in that framework: histories and transitions.

In the theory of branching time, the modal asymmetry between past and future finds its expression in a tree of moments that is linear toward the past and branches toward the future. Formally, a *branching time structure* is defined as a Kripke frame $\mathscr{M} = \langle M, < \rangle$ consisting of a non-empty set of moments M and an earlier-later relation $<$ on that set that fulfills the following three conditions: (1) there is no backward branching, (2) any two moments that are not comparable by the earlier-later relation $<$ share some common past up to a branching point, and (3) there is no last moment. An example of a branching time structure is provided in Fig. 1. The first condition captures the idea that, at any given moment, the past is fixed while the future may be open. It causes the structure to be tree-like. The second condition guarantees the unity of the structure. It ensures that we are provided with a single tree rather than with a forest and warrants the existence of branching points at the joints. The third condition, finally, reflects the idea that time does not end. There is always a future moment.

The definition of a branching time structure incorporates a global perspective on the tree of moments, viz. a perspective from outside of time. The tree of moments is described from a bird's eye view, as it were. If we circle above a branching time structure and look down, the tree of moments is lying there in its totality. The picture is an entirely static one. All moments are on a par. There is no distinguished actuality. There is no present, no past, and no future. There is only earlier and later, as witnessed by the order relation $<$ that unifies the tree of moments.

Naturally, the question arises what a branching time structure represents. One may be tempted to take the static tree of moments that

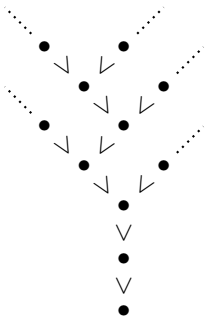


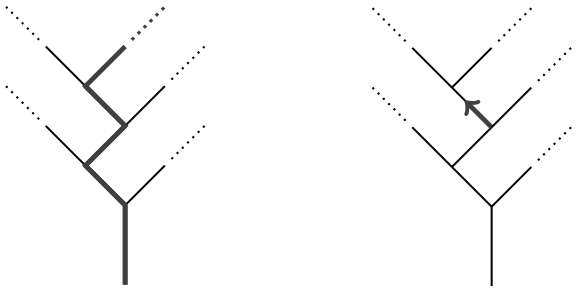
Figure 1: A branching time structure. The dots represent moments, and the earlier-later relation $<$ defines an order between those moments which takes the form of a tree.

we behold from above for an eternalist picture of the world or even of time itself. A branching time structure is not supposed to represent our world, however; nor does it represent time. Neither our world nor time branches. Arguably, both evolve linearly. What a branching time structure represents, when viewed from a bird's eye perspective, is the modal-temporal structure of an indeterministic world: it represents a world of possibilities.

Branching time structures harbor a wealth of possibilities. Owing to the existence of branching points, the earlier-later relation $<$ has a modal flavor, and it gives rise to possibilities, global and local ones, as illustrated in Fig. 2. Given a branching time structure $\mathscr{M} = \langle M, < \rangle$, we can identify maximal chains of moments that are linked by the earlier-later relation $<$. That is to say, we can carve out maximal $<$-linear subsets of M. Each such maximal $<$-chain in the tree of moments represents a *history*: a complete possible course of events. A history spans an entire possible temporal development of the world, and any two histories branch at some moment. They are historical alternatives.

Histories are possibilities, viz. *global* ones. But there are many more possibilities to be found in the modal-temporal structure of the world. If we zoom in on a branching point, we discover a plurality of *local* possibilities. Branching points are like forks in the road, where ways diverge. Each branching point allows for alternative possible future continuations, and each such immediate possible future continuation represents a local future possibility. Following Belnap (1999, [1]; 2005,

[2]), we refer to these local possibilities as *transitions*. One simple way to think about transitions—and, in fact, to draw them—is to view them as little arrows that each specify one possible direction at a branching point. A formal definition of a transition will be provided in section 5. As we shall see, the key to the transition semantics lies in the primacy of local over global possibilities.



(a) A history. (b) A transition.

Figure 2: Histories and transitions. (a) A history is a maximal $<$-chain of moments. (b) A transition is an arrow that specifies a possible direction at a branching point.

3 Actuality, Possibility, and Time

With these preliminaries in place, we now turn to the peculiar interrelation of actuality and possibility that we take to be at the heart of the theory of branching time. We have seen that in the static picture evoked by the definition of a branching time structure, there is no actuality yet. Actuality only enters the picture once we abandon our bird's eye view and take a local standpoint in the tree of moments, which is the perspective from which language is used. And as soon as we locate ourselves in the modal-temporal structure of the world, the purely static character of the structure disappears, and the tree gains dynamic traits.

So let us descend from our bird's eye perspective and settle on a moment in a branching time structure. By locating ourselves at a moment in the tree, we thereby mark that moment as present, and with a distinguished present, past and future enter the stage. At the same time, our

local standpoint partitions the tree into actuality, possibility, and counterfactuality, as illustrated in Fig. 3. Due to the absence of backward branching, the present moment determines a unique past, viz. the set of all moments preceding the present moment. The present and the past jointly provide a notion of actuality. They capture what has happened so far. But then, there may be various branches lying ahead of us. These branches represent possibilities for the future. None of them is actual yet; but any of them can be actualized. Branches that split off from the past course of events can best be dubbed counterfactualities. They are possibilities that once could have been actualized but actually were not.

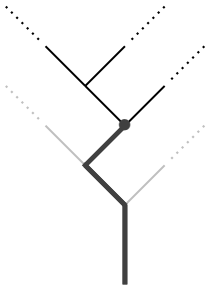


Figure 3: Actuality, possibility, and counterfactuality. The dot marks the present moment. The thick line represents actuality. Possibilities are indicated by thin black lines and counterfactualities by thin grey ones.

The notion of actuality that enters the picture once we take a local standpoint in the tree of moments is a temporal one, and it is dynamic. Actuality comprises only the present and the past. The future is not actual. The future is yet to come, and the branches spread out before us represent alternatives for the future: possibilities for actuality to unfold.

Note that the relation between actuality and possibility that we find in a branching time structure crucially differs from the one that we find in a possible worlds framework. In a possible worlds framework, our local standpoint in the pluriverse marks one of the possible worlds as the actual world, and the remaining possible worlds are considered mere modal alternatives to that actuality: they represent what could be the case, in some respect or another, but actually is not. The resulting picture is provided in Fig. 4. In the theory of branching time, in contrast, there is no actuality in the realm of possibilities, and the relation be-

tween actuality and possibility is a temporal rather than a modal one: the forking paths leading toward the future represent temporal alternatives *for* a dynamic actuality rather than modal alternatives *to* a given actuality.

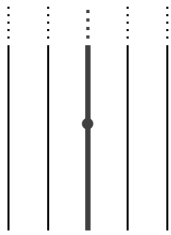


Figure 4: Actuality and possibility in a possible worlds framework. The dot specifies our local standpoint. The thick line represents actuality. Thin lines indicate possibilities.

The interrelation of actuality and possibility encoded in the tree of moments closely interacts with the notion of time, as illustrated in Fig. 5. When time passes, actuality evolves, and possibilities fade away into mere counterfactualities. The present moment becomes past, moments that have not been actual yet become actual, and possibilities that once have been open disappear. As actuality meanders through the maze of forking paths, at each branching point, one of the immediate future possibilities is actualized, ruling out the remainder. In the dynamic interrelation of actuality and possibility, the passage of time shines through. It presents itself in the constant loss of possibilities that accompanies the unfolding of actuality. In the end, all possibilities will have vanished, and there is only actuality. At present, however, there is no fact of the matter which of the forking paths actuality will tread.

4 Truth in the Absence of an Actual Future

The peculiar interrelation of actuality and possibility that we find in branching time structures poses a semantic challenge: how to assess whether a sentence about the future is true or false if there is no actual future? There are two traditional semantic approaches to branching time, which Prior (1967, [4]) refers to as *Peirceanism* and *Ockhamism*. In what follows, we briefly review how future truth is dealt with on those ac-

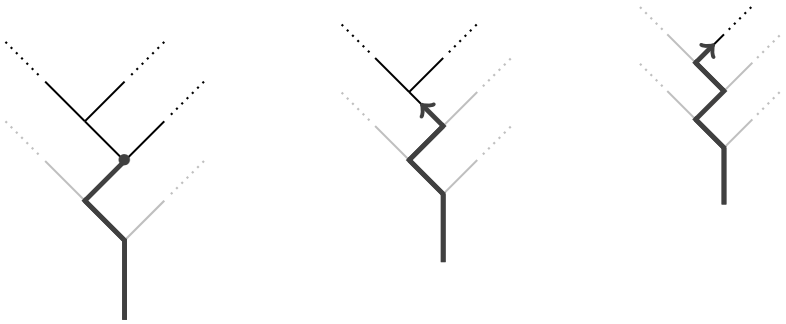


Figure 5: Actuality, possibility, and time. When time passes, at each branching point, one of the immediate future possibilities is actualized, and the remainder fades away.

counts and motivate the *transition semantics* proposed in Rumberg (2016, [5]).

In the Peircean semantics, sentences are evaluated from the local standpoint of a moment in the tree. Truth is relativized to a moment parameter. The local standpoint mirrors the perspective of language use, and it provides us with a notion of actuality. Yet, as we have seen, the relevant notion of actuality comprises only the present and the past. None of the forking paths leading toward the future is part of actuality. There is no actual future. There are only future possibilities, and, as of now, it is indeterminate which of them will be actualized. Due to the absence of an actual future, on the Peircean account with its local perspective, future truth comes out modalized. What will be the case is what is true in every possible future. Future truth and necessity are conflated.

In the Ockhamist semantics, on the other hand, in the absence of an actual future, a hypothetical one is introduced. The maze of forking paths is carved up into mutually exclusive, complete possible courses of events, viz. histories, and truth at a moment is relativized to one such overarching possibility. Sentences are evaluated at moment-history pairs. While the moment parameter pins down our local standpoint in the modal-temporal structure of the world, the history parameter provisionally singles out one complete possible evolution of actuality. What will be the case comes down to what is true in the given history. Modal

operators become interpretable as quantifiers over histories, and future truth and necessity come apart.

Let us briefly take stock here. In a way, Peirceanism and Ockhamism can be viewed as opposite extremes: they are strictly opposed to each other as regards their perspectives. Whereas on the Peircean account, truth is assessed solely from the local standpoint of a moment in the tree, the Ockhamist account makes use of a history as an auxiliary parameter of truth, and it thereby introduces a second, global perspective, viz. a perspective from the end of time. Despite—or rather, because of—the extreme, opposing perspectives, however, both Peirceanism and Ockhamism provide an entirely static picture of the interrelation of actuality, possibility, and time. On neither the Peircean nor the Ockhamist account is there room for actuality to unfold: the Ockhamist history parameter picks out a possible course of events that is already completely unfolded, and Peirceanism lacks a second parameter of truth.

The transition semantics is firmly rooted in our intuition of a dynamic actuality, which makes its way through the maze of forking paths. Just as in Ockhamism, truth at a moment is relativized to a possible course of events, a possible unfolding of actuality. However, possible courses of events are no longer identified with histories. Rather, possible courses of events are built up from local future possibilities: they are construed as sequences of transitions, which stretch linearly from the past toward a possibly open future. Incomplete possible courses of events become available as well, which can then be extended toward the future. The static Ockhamist history parameter is replaced by a dynamic transition parameter, which provides a second, local perspective in time that can be shifted independently of the moment parameter. The emerging picture is a tree of trees, as illustrated in Fig. 6. The possible future extensions of a sequence of transitions mirror the possible evolutions of actuality, and each possible future extension rules out at least one possibility.

The transition semantics generalizes both Peirceanism and Ockhamism and unifies both accounts in a single picture, as indicated in Fig. 6. By building on local future possibilities, it naturally bridges the gap between the local perspective inherent in Peirceanism and the global perspective characteristic of Ockhamism. Both the Peircean and the Ockhamist accounts turn out as limiting cases: they are to be found at

the extrema of the fine-grained picture evoked by the transition approach. Peirceanism amounts to the minimum of the tree of trees, which consists in the empty sequence of transitions. Ockhamism, on the other hand, corresponds to the maxima, which stand in for maximal sequences of transitions and, as such, are tantamount to histories.

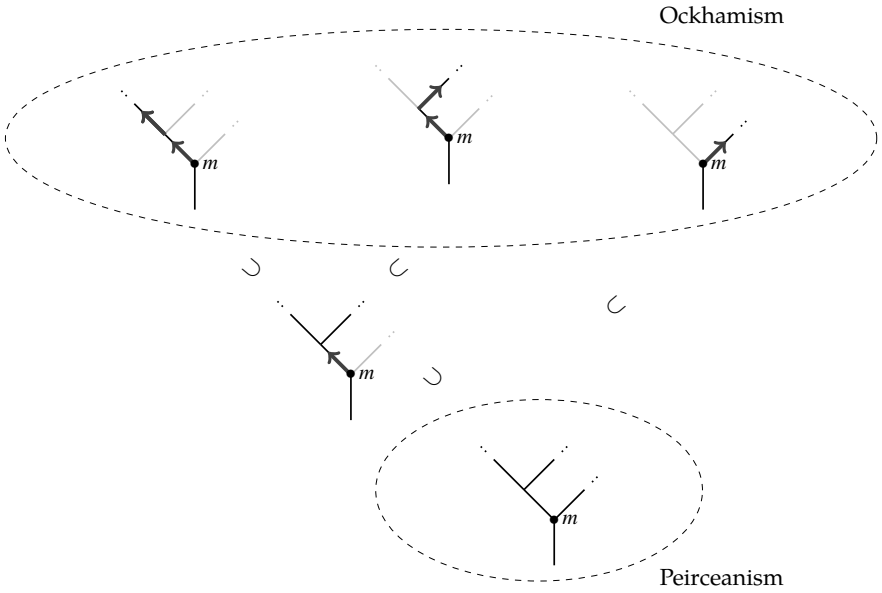


Figure 6: Possible courses of events in the transition semantics. Possible courses of events are construed as sequences of transitions. Peirceanism and Ockhamism are indicated as limiting cases.

5 The Maze of Forking Paths

The transition semantics draws on a fine-grained description of the maze of forking paths. The primacy of histories as global possibilities is renounced. Instead, possible courses of events are modeled by sequences of transitions. In this section, we show how the idea can be spelled out formally. We define the notion of a transition and illustrate how possible courses of events can be characterized by means of sets of transitions.

So far, we have only provided a pictorial representation of transitions.

We have depicted transitions as little arrows that specify a possible direction at a branching point, and we can find this idea back in the formal definition. A *transition* is defined as a pair consisting of a branching point m and a set of histories H that spans one of the alternative immediate possible future continuations of m. We call the branching point m the *initial* of the transition and the set of histories H its *outcome*. Conceived of as initial-outcome pairs, transitions are nothing but arrows: they are anchored in their initials, and they point toward their outcomes. The idea of an arrow is also visible in the notation: a transition with initial m and outcome H is written $\langle m \rightarrowtail H \rangle$.

Of course, it remains to be said which sets of histories span an immediate possible future continuation of a branching point and hence qualify as admissible outcomes of transitions. Here, we make use of the relation of *undividedness*, which captures the local branching behavior of histories at a given moment and allows us to treat discrete and continuous structures alike. At each branching point m, some pairs of histories branch while others continue to overlap for a while after m, in which case they are said to be undivided at m. The relation of undividedness at m is an equivalence relation on the set of histories containing the moment m and hence provides a partition of that set. Each cell of the partition captures one immediate possible future continuation of m, and for each cell, there is a transition with initial m whose outcome equals that set, as illustrated in Fig. 7.

As pointed out above, in the transition semantics, possible courses of events are construed as sequences of transitions. This creates the need to define an order $\prec$ between transitions which allows us to single out possible courses of events as linear sets of transitions; just as we have singled out histories as linear sets of moments in the order $<$. Consider two transitions $\langle m \rightarrowtail H \rangle$ and $\langle m' \rightarrowtail H' \rangle$. Obviously, for $\langle m \rightarrowtail H \rangle$ to precede $\langle m' \rightarrowtail H' \rangle$, it is necessary that the initials m and m' are ordered $m < m'$. But also the outcomes H and H' must stand in some suitable relation if the transitions are to be properly aligned. In particular, the outcomes must not be disjoint. Technically, this means that whenever $\langle m \rightarrowtail H \rangle$ precedes $\langle m' \rightarrowtail H' \rangle$, the outcomes H and H' must be ordered $H' \subset H$. Taking these two conditions together, we arrive at the following definition: we have $\langle m \rightarrowtail H \rangle \prec \langle m' \rightarrowtail H' \rangle$ if and only if ($m < m'$ and $H' \subset H$).

With the transition ordering $\prec$ at hand, we can now provide a formal characterization of possible courses of events in terms of sets of transi-

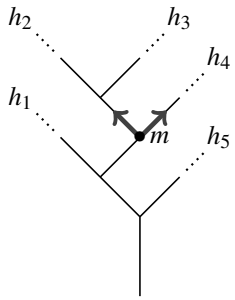


Figure 7: Transitions. The set of histories containing the branching point m has three elements: h_2, h_3, and h_4. The relation of undividedness at m partitions the set into two cells: $\{h_2, h_3\}$ and $\{h_4\}$. Hence, there are two transitions, $\langle m \rightarrowtail \{h_2, h_3\}\rangle$ and $\langle m \rightarrowtail \{h_4\}\rangle$, which are indicated by the two arrows.

tions. Analogously to the case of histories, we require that the relevant transition sets be linearly ordered by the relation $\prec$ between transitions; just as histories are linearly ordered by the relation $<$ between moments. Linearity provides us with a notion of modal consistency (cf. Müller 2014, [3]). In contradistinction to the case of histories, however, we do not require that the relevant transition sets be *maximal* linear sets. Rather, we only demand them to be *downward closed* in the transition ordering. Every $\prec$-chain of transitions that is closed toward the past depicts a possible course of events: a complete or an incomplete one. Maximal $\prec$-chains of transitions represent complete possible courses of events: they correspond to histories. Non-maximal $\prec$-chains of transitions, on the other hand, such as the one provided in Fig. 8, represent incomplete possible courses of events: they allow for alternative possible future continuations.

6 Future Truth, Contingency, and Stability

Unlike Peirceanism and Ockhamism, the transition semantics builds on a dynamic picture of the interrelation of actuality and possibility. Truth at a moment is relativized to a chain of transitions, which is closed toward the past but can be extended toward the future. In this section, we

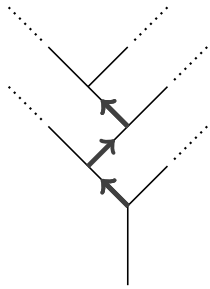


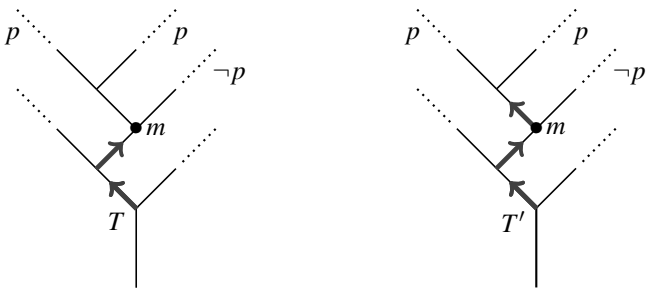
Figure 8: A set of transitions. The transition set contains three transitions, which are indicated by the three arrows. The transitions are linearly ordered by the relation $\prec$, and the set is downward closed. The possible course of events represented is an incomplete one, viz. one that allows for alternative possible future continuations.

illustrate what future truth amounts to in the transition framework, in the absence of an actual future.

In the transition semantics, sentences are evaluated at pairs m/T consisting of a moment m and a set of transitions T. The moment parameter m marks our local standpoint in the tree of moments, and the transition set T specifies a possible unfolding of actuality compatible with that moment. The language of the transition semantics extends the standard propositional language by a past and a future operator, P and F, a necessity operator $\Box$, and a so-called stability operator S, which is specific to the transition approach. The operators can roughly be divided into two classes: the temporal operators shift the moment parameter whereas the necessity and the stability operators vary the transition set.

In what follows, we will focus on the semantics of the future operator F and the stability operator S in order to illustrate how the contingency of the future is dealt with in the transition semantics. For this purpose, we consider the model provided in Fig. 9. The model contains a branching point m, which allows for two immediate possible future continuations: a p-continuation and a $\neg p$-continuation. We will be evaluating at the moment m with respect to two different transition sets: on the one hand, we consider the transition set T, which captures the actual past course of events up to the moment m; on the other hand, we consider the transition set T', which hypothetically specifies a possible

future extension of T.



(a) Evaluation w.r.t. m/T. (b) Evaluation w.r.t. m/T'.

Figure 9: Future truth, contingency, and stability. The sentence $\mathsf{F}p$ is false at the index m/T but true at the index m/T'. In particular, the sentence $\mathsf{F}p$ is contingent at the index m/T but stably-true at the index m/T'.

Let us first have a look at the future operator. The future operator F of the transition semantics has both Peircean and Ockhamist traits: the Peircean idea of universally quantifying over future possibilities is combined with the Ockhamist idea of relativizing truth to a possible course of events. What will be the case is what is true in every possible future that is admitted by the given transition set. That is, the transition parameter functions as a restrictor on the set of open future possibilities: all and only those future possibilities that are still open if actuality evolves as described by the given transition set are required to contain a witness for the future claim.

In order to get a feel for the restricting role of the transition parameter, let us see what happens if we evaluate the sentence $\mathsf{F}p$ in the model provided in Fig. 9 at the moment m with respect to the transition sets T and T', respectively. The transition set T does not rule out any of the future possibilities open at the moment m, and since there is one possible future continuation that lacks a witness for the future claim, the sentence $\mathsf{F}p$ cannot be true at the index m/T. If we wish to embrace bivalence, we are thus forced to say that the sentence is false at that index. Yet, as we shall see, the sentence comes out false at that index only contingently. By extending a transition set toward the future, we delimit the range of possible futures that we need to take into account. The pos-

sible future extension T' excludes the future continuation in which p fails to hold, rendering the sentence $\mathsf{F}p$ true at the index m/T'.

In the fine-grained picture underlying the transition approach, the truth value of a sentence about the future at a moment can change in the course of time, viz. under the possible future extensions of a given transition set, and the stability operator S constitutes a perspicuous means to capture this behavior. It provides us with a notion of relative settledness and brings to the fore the idea that contingencies about the future dissolve as time progresses. What is stably-true (stably-false) is what remains true (false) no matter how the future unfolds later on, i.e. no matter how we extend the given transition set. And what is neither stably-true nor stably-false is contingent.

With the stability operator at our disposal, we can express that the sentence $\mathsf{F}p$ is contingent at the index m/T: it is true at the moment m with respect to one possible future extension of T but false with respect to another. Relative to the future extension T', on the other hand, the truth value of the sentence stabilizes: the sentence $\mathsf{F}p$ is true at the index m/T', and it remains true at m under all possible future extensions of T'. It is worthwhile to note that in the transition semantics, the disjunction $\mathsf{F}\varphi \vee \mathsf{F}\neg\varphi$ is falsifiable. In our model, the sentence $\mathsf{F}p \vee \mathsf{F}\neg p$ is false at the index m/T. The disjunction can never be stably-false, however. Sooner or later, the contingency dissolves: $\neg\mathsf{S}\neg(\mathsf{F}\varphi \vee \mathsf{F}\neg\varphi)$ is a validity.

7 Conclusion

The interrelation of actuality and possibility that is at the heart of the theory of branching time is a rather peculiar one, and it differs crucially from the relation between actuality and possibility that we find in a possible worlds framework. In the theory of branching time, actuality is a temporal notion that comprises only the present and the past, and the branches leading toward the future represent possibilities for actuality to evolve: they are temporal alternatives *for* a dynamic actuality rather than modal alternatives *to* a given actuality.

The transition semantics is deeply rooted in the picture of a dynamic actuality. Possible courses of events are modeled by chains of transitions, which can be extended toward the future. In the absence of an overarching actuality which comprises an actual future, there is no need to conceive of possibilities as histories. In fact, transitions seem to be the

more natural choice. The local and dynamic nature of the transition approach provides an unprecedented view on the open future, on which future truth, contingency, and stability come apart.

The formal details are worked out in Rumberg (2016, [5]), where correspondence results are provided that expose Peirceanism and Ockhamism as limitations of the transition approach. In Rumberg and Zanardo (2018, [7]), the set-theoretic complexity of transition sets is dealt with, and axiomatizability results are established. The future is open for further research on transitions.

Acknowledgements

I would like to thank Patrick Blackburn for valuable comments and for his support. Many thanks also to Eric Raidl, Kurt Halter, Aron Grossmann, and Marco Gierga for their helpful feedback and for stimulating discussions. The paper was written while I was based at the University of Konstanz, and I acknowledge support by the German Research Foundation (DFG FOR 1614 TP7, MU1816/7-1).

Bibliography

[1] Belnap, N. (1999). Concrete transitions. In: Meggle, G. (ed.). *Actions, Norms, Values: Discussions with Georg Henrik von Wright*. Berlin: De Gruyter: pp. 227–236.

[2] Belnap, N. (2005). A theory of causation: *Causae causantes* (originating causes) as inus conditions in branching space-times. *The British Journal for the Philosophy of Science* **56**(2): 221–253.

[3] Müller, T. (2014). Alternatives to histories? Employing a local notion of modal consistency in branching theories. *Erkenntnis* **79**(3): 343–364.

[4] Prior, A.N. (1967). *Past, Present and Future*. Oxford: Oxford University Press.

[5] Rumberg, A. (2016). Transition semantics for branching time. *Journal of Logic, Language and Information* **25**(1): 77–108.

[6] Rumberg, A. (2016). *Transitions toward a Semantics for Real Possibility*. Quaestiones Infinitae: Publications of the Department of Philosophy and Religious Studies **98**. PhD thesis, Utrecht University.

[7] Rumberg, A. and Zanardo, A. (2018). First-order definability of transition structures. *Journal of Logic, Language and Information*. Online first: DOI 10.1007/s10849-018-9276-4.

TRL Semantics and Burgess' Formula

Roberto Ciuni
University of Padova, Italy
roberto.ciuni@unipd.it

Carlo Proietti
University of Lund, Sweden
University of Amsterdam, Netherlands
carlo.proietti@fil.lu.se

Abstract

ABE In this paper, we show that a particular thin red line semantics introduced by Bräuner, Øhrstrøm, and Hasle makes the so-called Burgess' formula $\Box G(\phi \rightarrow \Diamond F\phi) \rightarrow \Diamond G(\phi \rightarrow F\phi)$ invalid, and we discuss some reasons why the failure can be considered a problem for the semantics.

Keywords: Temporal logic, Branching time, Burgess' formula, Future contingents, Thin red line, Bundled Trees.

1 Introduction

The *Thin Red Line semantics* (TRL semantics for short) is a formal device that has been proposed to interpret future contingent propositions [4, 14].[1] This semantics extends the *trees* from [2, 15] with a *thin red line*

[1] Notice that TRL semantics has also been defended as an accurate formalization of William of Ockham's view on future contingents [13, 12]. As for *future contingents*, TRL is far from being the only formal approach. We will note some alternatives in the course of this paper.

function[2] $trl : T \longmapsto H_T$, which assigns a 'thin red line' $trl(t)$ to each moment $t \in T$—ideally, this is *the* actual history at t. In the TRL semantics sentences such as "tomorrow there will be a sea battle" (Fp) is true at t if and only if p is true at some later time t' in $trl(t)$.

The very idea of a 'thin red line' has met detailed criticism, notably by [1, 2] and [16]. In particular, criticism has been raised against the fact that $\phi \rightarrow PF\phi$ is not valid along the semantic account based on the so-called 'functional TRL' [16], which constitutes the standard version of TRL semantics. There are at least two reasons to consider this as an undesirable feature. First, $\phi \rightarrow PF\phi$ above expresses an appealing principle on the interaction between tenses, namely: if ϕ is true now, then it was the case that ϕ would be true (the so-called *retrogradation of truth*). Second, TRL semantics was meant to provide a conceptually sound account of future contingents, and this seems to have nothing to do with the principles that capture structural features of branching time. In a nutshell: choice of the TRL-based machinery should not conceal minimal temporal validities.

The TRL semantics by [4] restores the validity of $\phi \rightarrow PF\phi$ by introducing the notion of a set $C(t)$ of *counterfactual branches* along with the function trl. This is the set of histories that (1) pass through t, (2) are the thin red lines of some moment t' later than t. In this particular TRL semantics truth is relative to a moment in time t and a branch $h \in C(t)$. Given the truth conditions defined by [4], the introduction of $C(t)$ helps restore the validity of $\phi \rightarrow PF\phi$, as desired (see Section 2).

In this paper, we show that, although it fixes the problem above, the TRL semantics by [4] has a similar problem with the so-called Burgess' formula $\Box G(\phi \rightarrow \Diamond F\phi) \rightarrow \Diamond G(\phi \rightarrow F\phi)$ that has been discussed by [5] and [17]. In particular, [5] and [17] show that the formula is valid in *full trees*, but not in the so-called *bundled trees* (see Section 3). We prove our point in two steps. First, we exemplify a bundled tree $\mathscr{T}^{\mathscr{B}}$ that falsifies $\Box G(\phi \rightarrow \Diamond F\phi) \rightarrow \Diamond G(\phi \rightarrow F\phi)$ at a moment $t \in T$—this is the bundled tree from [17, §9]. Second, we prove that $\mathscr{T}^{\mathscr{B}}$ can be coherently extended with a trl function such that, for every $t \in T$, $C(t) \subseteq \mathscr{B}_t$—where $\mathscr{B}_t$ is the bundle of histories passing t. Since $\mathscr{B}_t$ guarantees the failure of $\Box G(\phi \rightarrow \Diamond F\phi) \rightarrow \Diamond G(\phi \rightarrow F\phi)$ in the initial bundled tree, we have that the failure is transferred in the corresponding bundled tree that

[2] This is also called *chronicle function* in [4].

includes the *trl* function. A consequence of this failure is that the TRL semantics by [4] falls short on the second methodological requirement we have presented above, which suggests that the semantics of the thin red line should not conceal the properties of the temporal structure.

The paper proceeds as follows. Section 2 introduces the basic temporal language and structures that are relevant for this paper, and introduces the TRL semantics from [4]. The section also discusses failure of $\phi \to PF\phi$ in the standard TRL semantics, its recovery in the semantics from [4], and it briefly describes the two methodological requirements above. Section 3 shows that Burgess' formula $\Box(G\phi \to \Diamond F\phi) \to \Diamond G(\phi \to F\phi)$ has a countermodel in the semantics from [4], and it also mentions two previously known failures. Section 4 discusses why failure of Burgess' formula is problematic for the proposal from [4]. Section 5 discusses some open issues and directions for future research. Finally, Section 6 presents some conclusions.

2 The thin red line: language and semantics

SYNTAX

The language $\mathscr{L}$ that we use in this paper is the standard option in philosophy [1, 10, 15], mathematical logic, [17, 21] and computer science [8, 9], when it comes to temporal reasoning.

Definition 1 (Language)**.** *Given a set $\mathscr{P}$ of atomic formulas, the language $\mathscr{L}$ is defined by the following BNF*:

$$\phi ::= p \mid \neg\phi \mid \phi \vee \psi \mid \phi \wedge \psi \mid \phi \to \psi \mid \Box\phi \mid P\phi \mid F\phi$$

where $p \in \mathscr{P}$, and the propositional connectives receive their standard interpretations. Formula $\Box\phi$ means "it is inevitably the case that ϕ", $P\phi$ means "it was the case that ϕ" and $F\phi$ means "it will be the case that ϕ". Duals of these expressions are $\Diamond\phi$, $H\phi$, and $G\phi$, respectively.

Remark 1. *This language dates back to* [*15, Ch. VII*]*. The reading of $F\phi$ is shared by nearly all formal approaches to branching-time logic and future contingents, and especially* Prior's Ockhamist branching-time logic*, and* TRL semantics*.* Supervaluationism *also shares this reading, but it equates*

the notion of truth (falsity) *with the notion of* necessary truth (falsity).[3] *By contrast, in* Prior's Peircean Branching-time logic $F\phi$ *reads 'Inevitably, it will be the case that* ϕ*', and it naturally generates a dual* $G\phi = \neg F \neg \phi$ *that reads 'Possibly, it will always be the case that* ϕ*'. In this logic, operator* G *is introduced as a primitive and* $G\phi$ *reads 'Inevitably, it will always be the case that* ϕ*'.*

TREES

Languages for indeterministic temporal reasoning are usually interpreted on *branching-time structures* known as *trees*, or some close variations of them [2, 15, 18]. In particular, trees are particular Kripke structures that are defined as follows:

Definition 2 (Trees)**.** *A tree* $\mathscr{T}$ *is a pair* $\langle T, < \rangle$ *where*:

- T is a set $\{t, t', t'', \dots\}$ of moments.
- $<$ is an *order – i.e. an irreflexive, asymmetric and transitive relation –* on T satisfying:

$$\begin{array}{ll} \forall t, t', t'' \in T : t' \leq t \text{ and } t'' \leq t & \\ \Rightarrow t' \leq t'' \text{ or } t'' \leq t' & (\textit{backward-linearity}) \end{array}$$

where $t \leq t'$ is short for '$t < t'$ or $t' = t$'. A crucial notion in branching-time semantics is that of a *history*. For every tree $\mathscr{T}$, a history is a maximal $<$-chain h of moments in T, i.e. it satisfies the following conditions:

$$\begin{array}{ll} t' \notin h \Rightarrow h \cup \{t'\} \notin H_{\mathscr{T}} & (\textit{maximality}) \\ \forall h \in H_{\mathscr{T}} : t \not\leq t' \text{ and } t' \not\leq t \Rightarrow t \notin h \text{ or } t' \notin h & (\textit{chain}) \end{array}$$

We denote with $H_{\mathscr{T}}$ the set of histories defined on tree $\mathscr{T}$. Furthermore, $H_t = \{h \in H_{\mathscr{T}} \mid t \in h\}$ is the set of histories that pass through the moment t. A tree is represented in Figure 1.

[3] More precisely, Supervaluationism take a formula like $F\phi$ to be *true* (false) if and only if for every history, ϕ holds good (fails to hold) at some later moment.

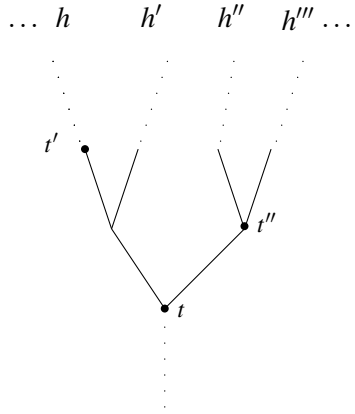


Figure 1: A tree

TRL SEMANTICS

The intuition behind TRL semantics is that, for every moment t, there is an actual history—'thin red line'—among the many passing through t. Views endorsing this intuition are usually referred to as *Actualism* [5], *True-futurism* [14], *Ockhamism* [3, 14], and they oppose so-called *Antactualism* [5], which is the view that no history passing through t can be singled out as the actual history of t. The most widespread version of TRL semantics define the thin red line as a function:

Definition 3. (Thin red line function) *For every tree $\mathscr{T}$, a thin red line (trl) function $trl : T \to H_{\mathscr{T}}$ is a function that satisfies the following two conditions:*

TRL1 $\forall t \in T : t \in trl(t)$

TRL2 $\forall t, t' \in T : t < t'$ and $t' \in trl(t) \;\Rightarrow\; trl(t') = trl(t)$

Condition TRL1 secures the 'effectivity' of trl: the thin red line is relevant for the continuation of events from t onwards. Condition TRL2 secures that the thin red line does not change until we remain on the same history.

The problem of Retrogradation of Truth. The standard TRL semantics evaluates sentences like $F\phi$ as true at t, relative to the function trl

$((t, trl) \models F\phi)$ if ϕ is true at a later moment t', relative to trl $((t', trl) \models \phi)$.[4] It has been noticed by many [3, 4, 14, 16] that standard TRL semantics makes $\phi \rightarrow PF\phi$ invalid. Figure 2 provides a countermodel. Here, we assume that there are no moments between t'' and t, that $trl(t'') = h'''$ and $trl(t) = h$ (the histories selected by trl are represented with thick lines), which is consistent with conditions TRL1 and TRL2. In the figure, we have $(t'', trl) \models p$ and $(t, trl) \not\models Fp$, since $trl(t) = h$ and $(t''', trl) \not\models p$ for every $t''' \in h$ later than t. Since, by assumption, there is no moment between t and t'', we have $(t'', trl) \not\models PFp$. This implies that $\not\models \phi \rightarrow PF\phi$.

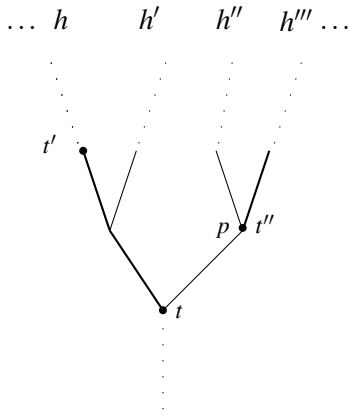


Figure 2: A counterexample to Retrogradation of Truth.

Failure of $\phi \rightarrow PF\phi$ is (nearly) universally considered a serious drawback of standard TRL semantics. Most notably, [16] argues that this failure shows the inadequacy of the 'thin red line' proposal, and even supporters of the proposal take the failure as an *indesideratum*. There is a good reason for this view: $\phi \rightarrow PF\phi$ expresses *within* the object language the semantical principle known as *retrogradation of truth*: *if it is true now that ϕ, then it was true that it would have been true that ϕ*.[5] This principle is ubiquitous in our linguistic practices involving tenses, as well as in social practices like betting or foreseeing. A theory of future contingents and temporal reasoning can hardly dispense with it. However, there is another reason for believing that failure of $\phi \rightarrow PF\phi$ is problematic. Such failure entails the invalidity of $\phi \rightarrow HF\phi$, which is an axiom

[4] We refer the reader to [3] for a detailed discussion on the semantics.

[5] This principle was also known as *de praesenti ad praeteritum* by Medieval logicians.

of *minimal temporal logic* and (partly) encodes a structural property of branching time, i.e. that *past* and *future* are mirror relations. Structural properties and their modally expressible counterparts (valid formulas) give us the principles of temporal reasoning. The role of a theory of future contingents (like TRL semantics) is altogether different: it is theory of truth-attribution to sentences (or assertions) about the future. Such a theory should be built on top of the structural properties of a branching-time structure (and of temporal reasoning) and should not conceal the properties of the temporal structure. But this is what standard TRL semantics is doing, by failure of $\phi \to PF\phi$.

AN ALTERNATIVE TRL SEMANTICS

[4] devises an alternative TRL semantics that validates $\phi \to PF\phi$. For every $t \in T$, this semantics takes into account also the *counterfactual thin red lines*, that is the histories that are thin red lines of some $t' \in T$ later than t ($t < t'$). The following definitions set the structures and notions needed for this.

Definition 4. **(Trl tree)** *A* trl tree $\mathscr{T}^{trl}$ *is a triple* $(T, <, trl)$ *where* $(T, <)$ *is a* tree *and* $trl : T \longmapsto H_{\mathscr{T}}$ *is a thin red line function.*

Definition 5. **(Trl model)** *A* trl model $\mathscr{M}^{trl}$ *is a pair* $(\mathscr{T}^{trl}, \upsilon)$ *where* $\mathscr{T}^{trl}$ *is a* trl tree *and* $\upsilon : \mathscr{P} \to 2^{T \times H_{\mathscr{T}}}$ *is a valuation function defined over the pairs* (t, h) *such that* $h \in H_t$*. For every* $p \in \mathscr{P}$*, the function satisfies:*

$$(t, h) \in \upsilon(p) \;\Leftrightarrow\; (t, h') \in \upsilon(p) \text{ for every } h' \in H_t$$

Definition 6. **(Counterfactual histories)** *For every trl tree* $(T, <, trl)$ *and moment* $t \in T$*, the set* $C(t)$ *of counterfactual histories at* t *is defined as follows:*

$$C(t) = \{h \in H_t \mid \forall t' \in h : t < t' \;\Rightarrow\; trl(t') = h\}$$

Temporal language is interpreted on the ground of the following satisfaction relation:

Definition 7. **(Satisfaction, 1)** *We interpret the formulas of* $\mathscr{L}$ *on trl models by the satisfaction relation* $\models$ *between trl models, moment-history pairs* (t, h)*—*

satisfying $h \in H_t$—and formulas. For all $h \in H_t$:[6]

$$\begin{array}{rcl}
\mathscr{M}^{trl}, (t,h) \models p & \Leftrightarrow & (t,h) \in \upsilon(p) \\
\mathscr{M}^{trl}, (t,h) \models P\phi & \Leftrightarrow & \exists t' < t : t' \in h \text{ and } \mathscr{M}^{trl}, (t',h) \models \phi \\
\mathscr{M}^{trl}, (t,h) \models F\phi & \Leftrightarrow & \exists t' > t : t' \in h \text{ and } \mathscr{M}^{trl}, (t',h) \models \phi \\
\mathscr{M}^{trl}, (t,h) \models \Box\phi & \Leftrightarrow & \forall h' \in C(t) : \ \mathscr{M}^{trl}, (t,h') \models \phi
\end{array}$$

Satisfiability in a model ($\mathscr{M}^{trl} \models \phi$) and trl-validity ($\models \phi$) are defined as usual. Condition $C(t) = \{h \in H_t \mid \forall t' \in h : t < t' \ \Rightarrow \ trl(t') = h\}$ for atomic sentences implies that $p \rightarrow \Box p$ for every $p \in \mathscr{P}$. It is easy to check that this extends to every formula ϕ that contains *no occurrence of* F. Now we prove that the semantics by [4] fixes the problem with *retrogradation of truth*.

Proposition 1. If the converse of $<$ is serial then $\models \phi \rightarrow PF\phi$ ($\phi \rightarrow PF\phi$ is trl-valid)

Proof. Suppose $\mathscr{M}^{trl}, (t,h) \models \phi$. Let us now analyze the truth conditions of $\mathscr{M}^{trl}, (t,h) \models PF\phi$ step by step. This equates with $\exists t' < t : t' \in h$ and $\mathscr{M}^{trl}, (t',h) \models F\phi$. From the backward linearity of $<$, $t' \in h$ for every $t' < t$, and since $H_t \subseteq H'_t$ for every $t' < t$, we have $h \in H'_t$. Since the converse of $<$ is serial, the existence of such a $t' < t$ is guaranteed. Left to check is whether $\mathscr{M}^{trl}, (t',h) \models F\phi$. This equates with $\exists t' > t : t' \in h$ and $\mathscr{M}^{trl}, (t',h) \models \phi$. Our initial moment of evaluation t guarantees this. Thus, $\mathscr{M}^{trl}, (t,h) \models PF\phi$. Since $\mathscr{M}^{trl}$ and (t,h) are arbitrary, this generalizes to any trl-model and pair (t,h) satisfying $h \in H_t$. As a consequence, $\phi \rightarrow PF\phi$ is trl-valid.

We briefly comment on the result.[7] The semantics by [4] avoids the problem that affects the traditional TRL semantics. Indeed, the history of evaluation $h \in H_t$ is fixed as the only relevant history for interpreting ϕ, $PF\phi$, and $F\phi$ and therefore $\phi \rightarrow PF\phi$ is valid. Just to get a concrete feeling of this, consider Figure 2. There, $(t'',h'') \models p$. Also, $(t,h'') \models Fp$, which in turn secures $(t'',h'') \models PF\phi$. Contrary to the standard TRL semantics, we do not need to switch to $h = trl(t)$ in order to evaluate $F\phi$ at t. Thus, the TRL semantics by [4] gets the *retrogradation of truth* back:

[6] We skip the clauses for the Boolean constructions, which are standard.

[7] Notice that that Proposition 1 and its proof also apply to Prior's Ockhamist semantics.

from the fact that it is true now that ϕ, we can infer that it was true that it would be the case that ϕ.

Relations with Prior's Ockhamist semantics. The TRL semantics by [4] can be seen as a variant of Prior's Ockhamist semantics introduced by [15]. The satisfaction relation of Prior's semantics is obtained from Definition 7 by replacing universal quantification on $C(t)$ with universal quantification of H_t in the truth clause of $\Box\phi$. This is a natural choice, since Prior's Ockhamist semantics does not define any *trl* function, or any device to single out *one* history as the only relevant history of evaluation (at a given moment).[8]

3 The failure of Burgess' formula

In this section we show that, although it fixes failure of $\phi \to PF\phi$, the TRL semantics from [4] runs into a similar problem with the so-called Burgess' formula [5]:

$$\Box G(\phi \to \Diamond F\phi) \to \Diamond G(\phi \to F\phi)$$

When it comes to Prior's Ockhamist semantics, this formula marks the difference between a number of different structures. In particular, given that semantics, the formula is valid on *full trees*, while it fails in *Kamp frames* [2, 17, 18], *Ockhamist frames* [2, 20], and bundled trees [2, 6, 17].[9] Here we show that the Burgess' formula is not valid in the TRL semantics from [4]. This is done by transforming a *bundled tree* which falsifies Burgess' formula [17, §9] into a corresponding TRL tree, where the failure is preserved. Let us define the notions of a *bundle* and of a *bundled tree*:

Definition 8. (**Bundle**) *Given a tree $\mathscr{T}$, a set $\mathscr{B}$ of histories is a* bundle *on $\mathscr{T}$ iff:*

[8] The TRL semantics by [4] has been criticized by [11] in the wake of its similarities with Prior's Ockhamist semantics. We do not go through the criticism here, and we refer the reader to [11] for it.

[9] Strictly speaking, the original formula appearing in [5] is not $\Box G(\phi \to \Diamond F\phi) \to \Diamond G(\phi \to F\phi)$, but $\Box G\Diamond F\phi \to \Diamond GF\phi$. These two formulas, however, are equivalent in all the families of structures we have mentioned. This is why the label 'Burgess' formula' is ambiguously used for both of them.

B1 $\mathscr{B} \subseteq H_{\mathscr{T}}$;

B2 $\forall t \in T : (\exists h \in \mathscr{B} : t \in h)$.

Definition 9. (Bundled Tree) *A bundled tree $\mathscr{T}^{\mathscr{B}}$ is a triple $(\mathscr{T}, \mathscr{B})$ where $\mathscr{T} = (T, <)$ is a tree and $\mathscr{B}$ is a bundle on $\mathscr{T}$*

A useful notation is $\mathscr{B}_t = \mathscr{B} \cap H_t$. The interpretation of formulas in a bundled tree works as for standard Prior's Ockhamist semantics, with the important exception that the history of evaluation is picked, for every $t \in T$, from $\mathscr{B}_t$ rather than from H_t. In particular:

Definition 10. (Model) *A* model *$\mathscr{M}^{\mathscr{B}}$ is a pair $(\mathscr{T}^{\mathscr{B}}, \pi)$ where $\mathscr{T}^{\mathscr{B}}$ is a* bundled tree *and $\pi : \mathscr{P} \to 2^{T \times H_{\mathscr{T}}}$ is a valuation function defined over all (t,h) such that $h \in \mathscr{B}_t$. For every $p \in \mathscr{P}$, the function satisfies:*

$$(t,h) \in \upsilon(p) \;\Leftrightarrow\; (t,h') \in \upsilon(p) \text{ for every } h' \in \mathscr{B}_t$$

Definition 11. (Satisfaction, 2) *We interpret the formulas of $\mathscr{L}$ on models by the satisfaction relation $\models$ between models, moment-history pairs (t,h)—satisfying $h \in \mathscr{B}_t$—and formulas. For all $h \in \mathscr{B}_t$:*

$$\begin{array}{rcl}
\mathscr{M}^{\mathscr{B}}, (t,h) \models p & \Leftrightarrow & (t,h) \in \pi(p) \\
\mathscr{M}^{\mathscr{B}}, (t,h) \models P\phi & \Leftrightarrow & \exists t' < t : t' \in h \text{ and } \mathscr{M}^{\mathscr{B}}, (t',h) \models \phi \\
\mathscr{M}^{\mathscr{B}}, (t,h) \models F\phi & \Leftrightarrow & \exists t' > t : t' \in h \text{ and } \mathscr{M}^{\mathscr{B}}, (t',h) \models \phi \\
\mathscr{M}^{\mathscr{B}}, (t,h) \models \Box\phi & \Leftrightarrow & \forall h' \in \mathscr{B}_t : \; \mathscr{M}^{\mathscr{B}}, (t,h') \models \phi
\end{array}$$

Satisfiability in a model ($\mathscr{M}^{\mathscr{B}} \models \phi$) and validity ($\models \phi$) are defined as usual. The truth clause for $\Box\phi$ implies that the operator $\Box$ quantifies only over the branches in $\mathscr{B}$, which may exclude some of the branches in H_t. This is in turn the key for the failure of the Burgess' formula on bundled trees, as is shown by the countermodel from [17, §9] that we present here. Consider the atomic formula $p \in \mathscr{P}$ and the model $\mathscr{M}^{\mathscr{B}} = (T, <, \mathscr{B}, \pi)$, where

- $T = \{(n,m) \in \mathbb{N} \times \mathbb{N} \mid 0 < n \leq m \text{ or } 0 = n = m\}$
- $(n,m) < (n',m') \;\Leftrightarrow\; (n = n' \text{ and } m < m') \text{ or } (n = m) < m'$
- $D = \{(n,m) \in T \mid n = m\}$
- $\mathscr{B} = H_{\mathscr{T}} \setminus \{D\}$

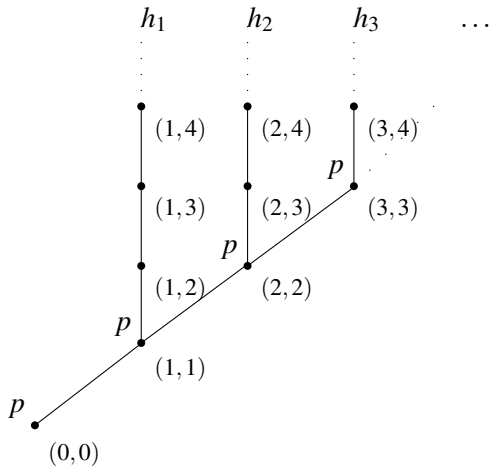


Figure 3: A countermodel to the Burgess' formula

- $(n,m) \in \pi(p) \Leftrightarrow (n,m) \in D$

The model is represented in Figure 3.

For the sake of simplicity, we name histories according to this convention: for every $n \in \mathbb{N}$ such that $n < \omega$, $h_n = \{(l,m) \in T \mid l = n$ and $l < m\} \cup \{(l,l) \mid l \leq n\}$. It is easy to check that $\mathscr{M}^{\mathscr{B}}, ((0,0),h_1) \models \Box G(p \rightarrow \Diamond Fp)$, while $\mathscr{M}^{\mathscr{B}}, ((0,0),h_1) \not\models \Diamond G(p \rightarrow Fp)$. Therefore Burgess' formula is falsified at $((0,0),h_1)$.

We now build a corresponding trl-model where the counterexample is preserved, to the effect that Burgess' formula is also invalid in the semantics from [4]. Given the model $\mathscr{M}^{\mathscr{B}}$, we define the trl model $\mathscr{M}^{trl} = (T', <', trl, \upsilon)$ where

- $T' = T$
- $(n,m) <' (n',m') \Leftrightarrow (n,m) < (n',m')$
- $trl(0,0) = \{(0,0)\} \cup \{(1,n) \mid n \in \mathbb{N}\}$

- $trl(n,m) = \{(n,m') \mid n < m'\} \cup \{(l,l) \mid l \leq n\}$

- $v(p) = \pi(p)$ for every $p \in \mathscr{P}$

It is straightforward to see that *trl* satisfies conditions TRL1 and TRL2 from Definition 3. Just to give a concrete grasp of the construction, we have[10] $trl(0,0) = trl(1,1) = trl(1,k) = h_1$ for $k \geq 1$, $trl(2,k) = h_2$ for $k \geq 2$, $trl(3,k) = h_3$ for $k \geq 3$ and so on. Also, notice that the history $D = \{(n,m) \in T \mid n = m\}$ is the trl of *no* moment in T': $trl(t) \neq D$ for every $t \in T$.

We now show that falsification of Burgess' formula at $(0,0)$ is transferred to $\mathscr{M}^{trl}$. Given Definition 6 and the construction above, it is straightforward that $C(0,0) = \{h_1\}$ and for every $(n,n) \in D$, if $0 < n$, then $C(n,n) = \{h_n, h_{n+1}\}$. We have $\mathscr{M}^{trl}, ((0,0),h_1) \models \Box G(p \to \Diamond Fp)$. In order to see this, consider that (i) $\mathscr{M}^{trl}, ((0,0),h_1) \models G(p \to \Diamond Fp)$. Indeed, for every $(1,k)$ with $k > 1$, $\mathscr{M}^{trl}, ((1,k),h_1) \models p \to \Diamond Fp$ trivially holds since $\mathscr{M}^{trl}, ((1,k),h_1) \not\models p$. As for $(1,1)$, we have $\mathscr{M}^{trl}, ((1,1),h_2) \models Fp$. Since $h_2 \in C(1,1)$, this implies (ii) $\mathscr{M}^{trl}, ((1,1),h_1) \models p \to \Diamond Fp$. From (i) and (ii), $\mathscr{M}^{trl}, ((0,0),h_1) \models G(p \to \Diamond Fp)$ follows. Since $C(0,0) = \{h_1\}$, this suffices to have $\mathscr{M}^{trl}, ((0,0),h_1) \models \Box G(p \to \Diamond Fp)$. At the same time, however, we have $\mathscr{M}^{trl}, ((0,0),h_1) \not\models \Diamond G(p \to Fp)$. Indeed, $\mathscr{M}^{trl}, ((0,0),h_1) \not\models G(p \to Fp)$ since $\mathscr{M}^{trl}, ((1,1),h_1) \not\models p \to Fp$: p is true at $((1,1),h_1)$ and false at any pair $((1,k),h_1)$ for $k > 1$. Again, Burgess' formula is falsified at $((0,0),h_1)$.

Formal differences aside, the falsification of Burgess' formula on the model $\mathscr{M}^{\mathscr{B}}$ and the trl model $\mathscr{M}^{trl}$ above relies on the very same fact: exclusion of D. Indeed, suppose we included D in the set of histories that are relevant for evaluating formulas of the form $\Box\phi$ and $\Diamond\phi$ at any $(n,m) \in T$ such that $n = m$. This yields $((n,m),D) \models G(p \to Fp)$. More: D is the only history relative to which the formula is satisfied. Thus, the exclusion of D from $\mathscr{B}$ is a crucial move in the counterexample by [17]. As for our trl-model $\mathscr{M}^{trl}$, a consequence of the construction above is that $D \notin C(n,m)$ for every $(n,m) \in T$. This is secured by the fact (that is easily checked) that $C(t) \subseteq \mathscr{B}_t$ for every $t \in T$. But once again, D is the only history relative to which the formula is satisfied.

[10] In what follows, we will abuse notation a bit and write $trl(n,m)$ rather then $trl((n,m))$, for the sake of readability. The same convention applies to C.

Burgess' formula is not the only principle of temporal reasoning that fails in the semantics from [4]. As [4, 3, 14] notice, another notable failure is $F\Diamond F\phi \rightarrow \Diamond FF\phi$, which connects the futurity of a possibility with a possibility.[11] In order to see this, change the trl-model that we constructed above (see Figure 3) by just imposing $(1,k) \notin \upsilon(p)$ for every $k \geq 1$ and keeping $(2,2) \in \upsilon(p)$. From this and $C(1,1) = \{h_1, h_2\}$, we have $((1,1),h_1) \models \Diamond F\phi$. From this and $C(0,0) = \{h_1\}$, we have $((0,0),h_1) \models F\Diamond F\phi$. However, from $(1,k) \notin \pi(p)$ for every $k \geq 1$ and $C(0,0) = \{h_1\}$, we have $((0,0),h_1) \not\models \Diamond FF\phi$.

Also, $H(G\phi \rightarrow \Diamond G\phi)$ is not valid. Take again the model depicted in Figure 3, and let us include history D as a relevant parameter of evaluation. We have $((0,0),D) \models Gp$. But since $C(0,0) = \{h_1\}$, we have $((0,0),D) \not\models \Diamond Gp$. This in turn implies that $((2,2),D) \not\models H(Gp \rightarrow \Diamond Gp)$.

Contrary to Burgess' formula, these principles are valid, relative to Prior's Ockhamist semantics, on both trees and bundled trees. Both formulas seem to capture principles that are rooted in our intuitions about time: $F\Diamond F\phi \rightarrow \Diamond FF\phi$ implies that *future possibilities just do not emerge anew*; $H(G\phi \rightarrow \Diamond G\phi)$ implies that what will simply happen by chance is a ground for what future possibilities are. Notice that failure of $H(G\phi \rightarrow \Diamond G\phi)$ (but not of $F\Diamond F\phi \rightarrow \Diamond FF\phi$) is fixed by a different TRL semantics that is introduced in [3]. Due to limits of space, we cannot discuss this further proposal here, but we plan to consider this semantics in future research.

4 Discussion

Failure of Burgess' formula implies that the TRL semantics from [4] does not meet one of the methodological requirements that we have outlined in Section 2: this particular theory of future contingents conceals some structural property of temporal reasoning—in our case, a principle expressing some specific interactions between tense and modality. We briefly discuss Burgess' formula and its place in branching-time logic, in order to get a better idea of why its failure is a drawback for the semantics from [4].

[11] Here, by 'a possibility' we mean a fact that is expressed by a true formula $\Diamond F\phi$.

The Burgess' formula states that if it is inevitable that a fact will always keep being a future possibility, then it is possible that it will alway keep happening from time to time, at least if it happens sometimes. This interaction between tense and modality is valid if, for every moment $t \in T$, there is at least one history $h \in H_t$ that can be associated with a limit ordinal. This condition is in turn secured by the (full) trees that we have presented in Section 2, but it is failed by some bundled trees (like the one from [17, §9] and Section 3). Thus, Burgess' formula describes a property of the temporal structures that [4] presupposes for the interpretation of temporal language (that is, full trees). In Section 2, we suggested that validity of this kind of formulas (and hence, of properties) should not be concealed by our theories of future contingents, since these aim at providing an adequate mechanism of truth attribution to some sentences, not to tell us what properties our temporal structures enjoy. In light of failure of $\Box G(\phi \to \Diamond F\phi) \to \Diamond G(\phi \to F\phi)$, we can conclude that the TRL semantics from [4] really clashes with the methodological requirement above.

Besides, many temporal logicians regard Burgess' formula as a desirable principle of temporal reasoning, to the effect that the semantics from [4] does not meet the first methodological requirement from Section 2 either: it invalidates a widely accepted principle of temporal reasoning. In particular, a discussion on the descriptive adequacy of bundled trees (and then, on the status of Burgess' formula) is carried by [2, §7A.6], [12, pp. 268–269], and [18, pp. 222–223]. The discussion focuses on examples such as the following:[12]

($\star$) Inevitably it will be the case that, if radium atom a has not yet decayed (p), then a might not decay in the next moment but it might also decay in the next moment.

($\star\star$) Inevitably, it will be the case that the atom a has not yet decayed and will decay on the next moment.

In the language that we deploy in this paper, ($\star$) is expressed by $\Box G(p \to (\Diamond Fp \wedge \Diamond F\neg p))$; ($\star\star$) is expressed by $\Box F(p \wedge G\neg p)$.[13] The first implies the antecedent of the Burgess' formula, the second denies its consequent.

[12] The example is an adjustment of an example from [2].

[13] Of course, we are assuming that $\neg p \to G\neg p$: atom a stays decayed after he gets decayed.

[2] and [18] maintain that ($\star$) and ($\star\star$) are mutually incompatible. In turn, this implies that they support validity of Burgess' formula and the adequacy of full trees as temporal structures. By contrast, [12] admits for the compatibility of ($\star$) and ($\star\star$), thus implying that bundled trees (or other related structures) may provide an adequate representation of time.

The view from [2, 18] has raised a wider consensus than the view from [12], to the effect that many see full trees as the adequate structures for interpreting branching-time logic. We do not take part in the debate here, although we believe that bundled trees deserve more consideration.[14] That being said, we believe that the issue concerning the validity of Burgess' formula and use of full trees (or bundled trees) should be settled by arguing about the desirability of the temporal property it expresses.[15] By contrast, the semantics from [4] decides for invalidity of Burgess' formula by a machinery that is justified by considerations on future contingents. In turn, this has nothing to do with the properties we wish (or wish not) ascribe to our temporal structures. Hence, we believe that appeal to the requirements suggested in Section 2 is cogent in our case, and that it justifies the claim that failure of Burgess' formula is a drawback of the TRL semantics from [4].

Similar considerations apply to the failures of $F\Diamond F\phi \rightarrow \Diamond FF\phi$ and $H(G\phi \rightarrow \Diamond G\phi)$, with the difference that the two principles seem to be deeply rooted in our intuitions as $\phi \rightarrow PF\phi$ is. Thus, their failure is equally unsatisfactory.

5 Open issues and perspectives for future research

The failure discussed in this paper raises a number of questions. Of course, a natural question is whether the TRL semantics from [4] can be adjusted in order to restore validity of the Burgess' formula, without

[14] A recent addition to the debate is [7], which introduces a particular game-theoretical scenario based on an infinitely repeated Prisoner Dilemma. This scenario cannot be modeled on full trees.

[15] To this concern, we find that the considerations put on the table from [2, 18] are far from conclusive. In particular, we agree with [21] that talk of incompatibility (or compatibility) actually *presupposes* the choice of full trees over bundled trees (or vice versa), and thus it does not help settle the issue.

at the same time renounce to provide a theory of future contingents. A similar question concerns the failures of $F\Diamond F\phi \rightarrow \Diamond FF\phi$ and $H(G\phi \rightarrow \Diamond G\phi)$.

We believe that a real understanding of this issue can be provided only if we get a clear and rigorous way of what is the import of the above failures for the temporal structures we use in our semantics. We have already mentioned that the so-called *limit closure* is lost if Burgess' formula is not valid. However, it would be interesting to explore which properties are called off by the other two failures.

In our opinion, *correspondence theory* [19] is the most promising way to approach the problem. This implies shifting from the second-order theory of (full) trees to a first-order theory of branching time. We know from [17, 20] that this is no problem: temporal structures such as Ockhamist frames [20] provide first-order theories of time; moreover, the notion of validity in such structures is equivalent to the notion of validity in bundled trees. In future research, we wish to deploy correspondence theory in order to connect (failure of) the three formulas above with (lack of) first-order properties of suitable temporal structures. An assessment of the cogency of these properties (or the admissibility of their lack) will in turn provide a rigorous way to understand whether the TRL theorist in the style of [4] can find a way out of the failure—or justify them with a different view on time—or must give up on the tenability of the solution that she is proposing.

Another interesting open issue is whether failure of Burgess' formula affects also the semantics from [3]. We plan to explore the topic in a future paper.

6 Conclusion

In this paper, we have introduced the TRL semantics from [4] (Section 2), we have shown that it fails Burgess' formula $\Box G(\phi \rightarrow \Diamond F\phi) \rightarrow \Diamond(G \rightarrow F\phi)$ (Section 3), and we have discussed the conceptual import of this failure (Section 4). The semantics from [4] has been devised in order to fix another failure of the standard TRL semantics, namely that of $\phi \rightarrow PF\phi$. In section 2, we comment on this failure and suggest that a formal theory of future contingents should not conceal the validity of principles of temporal reasoning. We prove failure of Burgess' formula

in Section 3. In doing this, we rely on the fact that the formula is not valid on the so-called *bundled trees*. In particular, we consider a particular *bundled tree* from [17] that provides a counterexample to Burgess' formula, we turn it into a corresponding *trl tree* in the style of [4], and we show that failure of Burgess' formula is preserved by the construction. We also briefly comment on two other failures of standard temporal principles in the semantics from [4]. In Section 4, we discuss the failure and explain why it proves unfortunate: beside violating the methodological requirement outlined in Section 2, this failure makes the semantics of [4] take side in a debate that has no apparent connection with the issue of future contingents. Finally, we discuss some open issues and perspectives for future research.

Acknowledgements

Work of Carlo Proietti was supported by the European Commission under the Marie Skłodowska-Curie Action H2020-MSCA-IF-2016 (748421) and by the Sveriges Riksbanken (RJ: P16-0596:1).

Bibliography

[1] Belnap, N. and Green, M. (1994). Indeterminism and the thin red line. *Philosophical Perspectives* **8**: 365–388.

[2] Belnap, N., Perloff, M. and Xu, M. (2001). *Facing the Future. Agents and choices in our Indeterminist World*. Oxford: Oxford University Press.

[3] Braüner, T., Hasle, P. and Øhrstrom, P. (1998). Ockhamistic logics and true futures of counterfactual moments. In: Kathib, L. and Morris, R. (eds.): *International symposium on temporal representation and reasoning* (*TIME-98*): 132–139.

[4] Braüner, T., Hasle, P. and Øhrstrom, P. (2000). Determinism and the origins of temporal logic. In: Barringer, H., Fisher, M., Gabbay, D. and Gough, G. (eds.): *Advances in Temporal Logic. Applied Logic Series, volume 16*. Dordrecht: Springer: pp. 185–206.

[5] Burgess, J.P. (1978). The unreal future. *Theoria* **44**(3):157–179.

[6] Burgess, J.P. (1979). Logic and time. *Journal of Symbolic Logic* **44**: 556–582.

[7] Ciuni, R. and Lorini, E. (2017). Comparing semantics for temporal STIT logic. *Logique et Analyse*, online first.

[8] Emerson, E.A. and Halpern, J.Y. (1986). "Sometimes" and "Not Never" Revisited: On Branching versus Linear Time Temporal Logic. *Journal of the Association for Computing Machinery* **33**(1): 151–178.

[9] Emerson, E.A. and Sistla, A.P. (1984). Deciding full branching-time logic. *Information and Control* **61**(3): 175–201.

[10] MacFarlane, J. (2003). Future Contingents and Relative Rruth. *The Philosophical Quarterly* **53**: 321–336.

[11] Malpass, A. and Wawer, J. (2012). A future for the thin red line. *Synthese* **188**: 117–142.

[12] Øhrstøm, P. and Hasle, P. (1995). *Temporal Logic: From Ancient Ideas to Artificial Intelligence*. Berlin: Springer Verlag.

[13] Øhrstrøm, P. (1984). Anselm, Ockham and Leibniz on divine foreknowledge and human freedom. *Erkenntnis* **21**(2): 209–222.

[14] Øhrstrøm, P. (2009). In Defence of the Thin Red Line: A Case for Ockhamism. *Humana.Mente* **8**: 17–32.

[15] Prior, A.N. (1967). *Past, Present and Future*. Oxford: Oxford University Press.

[16] Restall, G. (2011). Molinism and the thin red line. In: Perszyk, K. (ed.): *Molinism: The Contemporary Debate*. Oxford: Oxford University Press: pp. 227–238.

[17] Reynolds, M. (2002). Axioms for Branching Time. *Journal of Logic and Computation* **12**(4): 679–697.

[18] Thomason, R. H. (2002). Combinations of tense and modality. In: Gabbay, D. and Guenthner, F. (eds.): *Handbook of Philosophical Logic*, volume 7, 2nd edition. Berlin: Springer Verlag: pp. 205–233.

[19] van Benthem, J. (2002). Correspondence theory. In: Gabbay, D. and Guenthner, F. (eds.): *Handbook of Philosophical Logic*, volume 3, 2nd edition. Berlin: Springer Verlag: pp. 325–408.

[20] Zanardo, A. (1996). Branching-time logic with quantification over branches: The point of view of modal logic. *Journal of Symbolic Logic* **61**(1): 1–40.

[21] Zanardo, A. (2006) Quantification over Sets of Possible Worlds in Branching-Time Semantics. *Studia Logica* **82**(3): 379–400.

The Role of Time in Phronetic Activities

Anne Gerdes
University of Southern, Denmark
gerdes@sdu.dk

Abstract

When we are engaged in phronetic activities, the notion of time plays a crucial role in our exercise of judgment in specific cases. For instance, undertaking an obligation implies that we seek to figure out what we ought to do in a given situation by considering plausible future scenarios with morally preferably outcomes on the background of the current situation. Likewise, in deliberating about whether we did right or wrong, we evaluate past actions by zooming in on some specific past circumstances, which are chosen as points of departure for judging a given past course of events. Also, judgments about situations involving moral wrongdoing and accountability often involve reasoning from counterfactual scenarios based upon revisiting an alternative past version of future courses of events, which might preferably have taken place. When we act morally this limits our future space of freedom of action and thereby involves a certain amount of risk-taking. The learning process inherent in such risk-taking activities drives the kind of engagement needed to cultivate us to become phronimos. In continuation thereof, it is argued that an artificial ethical agent shall presumably never be phronimos in the strong sense as it would lack the ability to experience and learn from the relation between time and risk and the role this relation plays in moral life.

Keywords: time, phronesis, artificial moral agents, branching time models, machine learning.

1 Introduction

This paper discusses the role of time in relation to phronetic activities, i.e. activities focusing on situated deliberation about moral choices followed by the enactment of selected moral actions. The notion of choice may be clarified with reference to Aristotle's notion of choice as a voluntary action, which we decide on by previous deliberation - "Choice will be a deliberate desire of things in our power [...]" (Aristotle 1934: NE, Book III. iii. 17-iv. 3, p. 141, [2]). This means that we do not deliberate about the present moment since things that happens instantaneously, at this particular moment, cannot be deliberated upon. Also, the past is necessary now and it is not in our power to make choices that may change what happened in the past. Hence, deliberation is restricted to future alternatives. However, we often reflect on past events in a judgmental context (Thomason 1981, [17]), when engaging in after the fact evaluations of past events. Here, we deliberate about the past by applying counterfactual reasoning and move back to a point in time in which alternatives were still open to us, i.e., to the point in time in which 'ought' still implied 'can' and would have allowed us to act differently (Thomason 1981, p. 174, [17]).

To further refine the above-mentioned observations concerning this fundamental asymmetry between the past and the future, i.e., the openess of the future and the nececesity of history, the second part of the paper introduces Prior's philosophy of time (1949). His preliminary outline of a tensed ontology takes as its point of departure presentism, i.e. the notion of the dynamic aspects of time as changing, which is reflected in the interplay between the present, which is real, the past, which was real, and the future, which will become real (Øhrstrøm, Schärfe and Ploug 2010, [23]). Here, his development of a notion of branching time makes it possible to settle the truth value of future contingencies, while avoiding determinism. In that sense, Prior's tense logic conceptualizes the importance of the relation between time, free choice, and ethics.

Based on this outline of the philosophy of conceptual structures of time, the third part of the paper introduces the notion of phronesis, i.e., the kind of practical wisdom or experience-based knowledge that is crucial for navigating in social situations to a decision, making moral choices with proficiency. Here, the kairos of such phronetic activities

implies that the phronimos can judge the incentives as well as the stakes involved when selecting the right moment to act.

Finally, in part four, the paper elaborates on the kairos of phronetic activities. It is argued that an artificial ethical agent shall presumably never be phronimos, as it would carry out ethical reasoning with no real sense of how time affects our moral life. That is excatly why, when approaching a future in which an artificial ethical agent will be able to make ethical decisions, we should ensure that such a system will be capable of presenting us with proper ethical evaluations of its moral choices.

2 Time, future contingents, and free choice

Prior's temporal realism resonates well with Aristotle's notion of free choice as the asymmetric structure of time reflects the relation between time and free choice. The quotation below represents some of Prior's early thoughts on tense logic as reflected in the paper *Some Free Thinking about Time*, which laid the foundation for his break through contribution *Time and Modality* (1957, [13]), which was followed up by his milestone book *Past, Present, and Future* (1967, [15]).

> One of the big differences between the past and the future is that once something has become past, it is, as it were, out of our reach - once a thing has happened, nothing we can do can make it not to have happened. But, the future is to some extent, even though it is only to a very small extent, something we can make for ourselves. And this is a distinction which a tenseless logic is unable to express. In my own logic with tenses I would express it this way: We can lay it down as a law that whatever now is the case will always have been the case; but we can't interchange past and future here and lay it down that whatever now is the case has always been going to be the case – I don't think that's a logical law at all; for if something is the work of a free agent, then it wasn't going to be the case until that agent decided that it was. But if happenings are just properties timelessly attached to dates, I don't see how you can make this distinction.
>
> (Prior 1958, p. 2, [14])

Prior emphasizes the importance of a so-called A-theoretical representation of time, which is in accordance with our everyday notion of time (Prior 1967, p. 2, [15]). Here, we have an arrangement of events ordered by A-concepts, 'past', 'present', and 'future', which mirrors the asymmetrical relation between the past (as necessary in the present) and the future (as possible in the present). Consequently, an A-theoretical description captures the dynamics of change as time passes by - as opposed to the static, tenseless symmetrically linear ordered B-time-series, by means of which time positions are presented in a fixed temporal order. Hence, in what follows, the A-theoretical idea of the asymmetry of time is presented by introducing the notion of branching time, which captures not only how future developments of events can be seen as either possible or necessary, but also how the role of time is important to account for moral choice.

In the context of time and ethics, it is striking, that despite the fact that Prior and von Wright exchanged ideas, we see no traces of tense logic in Von Wright's work on deontic logic, although he acknowledged Prior's work on tense logic –- "What I have seen of your work, however, makes it clear that it is important. [...] It must be very satisfying to you to know that you started this new and exciting branch of logical study. [...] I'm sure it will have a great future" (Jakobsen, Øhrstrøm and Hasle. 2017, p. 11, [7]). Also, before *Past, Present and Future*, Prior worked on deontic logic, and in the paper: *The Paradoxes of Derived Obligation* (Prior 1954, [12]), he discussed von Wright's work, and again, von Wright commented on the critique (Wright 1956, [19]). However, Prior never fullfilled the ambition he had concerning the incorporation of deontic and tense logic. Nevertheles, Prior's ideas on branching time is highly relevant for a conceptualization of the interrelatedness of time and ethics. An example of Prior's early ideas on branching time is reflected in the paper: *Postulate-sets for Tense-logic*, which can be found in the Prior Nachlass archive[1].

> In these models the course of time [...] is represented by a line which, as it moves from left to right (past to future), continually divides into branches, so that from any given point on the diagram there is a unique route backward (to the left; to the past) but a variety of routes forwards (to the right; to

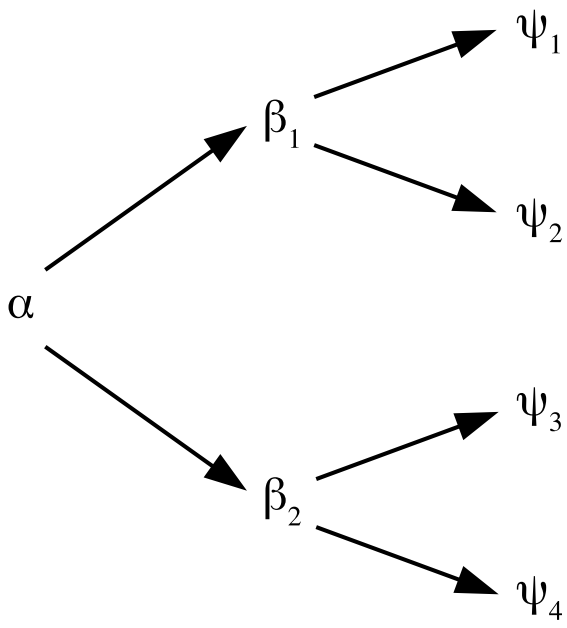
[1]The Nachlas of A.N. Prior: http://nachlass.prior.aau.dk/

> the future). In each model there is a single designated point, representing the actual present moment; and in an Occamist model there is a single designated line (taking one only of the possible forward routes at each fork), which might be picked out in red, representing the actual course of events.
>
> (Prior 1966, p. 5, [11])

Subsequently, one of the major topics in his influential *Past, Present and Future* was the discussion of issues concerning divine foreknowledge and free choice, presented in his investigations of puzzles concerning the relation between the truth-value of future states of events and indeterminism. Here, Prior presented the notion of branching time as a useful framework for discussing determinism and indeterminism.

The branching time model below illustrates four so-called chronicles, each of which represents possible future courses of events, only one of which will be realized in the future. This means that the unrealized chronicles reflect counterfactual courses of events (what would have been realized given counterfactual circumstances). Different models have been suggested to account for future contingents. Hence, in *De Interpretatione, Chapter IX*, Aristotle's example of "the sea-battle tomorrow" discusses whether statements about the future ("there will be a sea battle tomorrow", "there will not be a sea battle tomorrow") can be said to be true today. Here, Aristotle claimed that the future is indeterministic, i.e., propositions about future events are neither true nor false (Anscombe 1967, [1]). However, in contemporary research concerning an interpretation of the idea of the true future, work has been dedicated to the development of a satisfactory representation of a future operator in branching time models (Thomason 1981, [17]; Øhrstrøm 2009, [22]). Prior maintained that indeterministic tense-logic is possible without rejection of the principle of bivalence. He gave two interpretations of the idea of branching time, the so-called "Ockhamist system" and the "Peircian System", one its fragment (Prior 1967 p. 130, [15]). According to the Ockhamist system, from a God-eyes-perspective, we may say that there exists one possible future, which is true now. This also reflects our intuition about time as we in the present act in accordance with what we expect to happen in the future, even though it is not possible for us to know whether the expected future eventuates or not. Clearly, as has been shown by Øhrstrøm and Hasle (2011, [24]) to the Ockhamist,

there is a distinction between three expressions of future, viz. 'it will be the case that p', 'it will possibly be the case that p', and 'it will necessarily be the case that p', which may be formalized by introducing modal operators. Hence, we may at time α (the root time shown in the diagram) ascribe *prima facia* truth value relative to a designated route in the branching time model, implying that the truth-value of future contingencies is relative to a specifically selected history. In this manner, it becomes possible to distinguish between necessary and possible future courses of events, allowing one to assign truth-values to future contingencies while avoiding determinism.



Of the fragment that Prior called the Peircean system, Prior remarks that:

> The Peircean system can only say 'it will be that *p*' when *p*'s futurition is necessary, when it is not necessary but will occur all the same, he has to say that 'It will be that *p*' is false;

> the sense in which it is true eludes him. But to the Peircean, the Ockhamist seems to treat what is still future in a way in which it would only be proper to treat what *has been future* – he views it as it would be proper to view it from the end of time. (Prior 1967, p. 130, [15])

Prior favored the Peircean system since he found that it did not make sense to refer to a true future now with reference to a *prima facia* future, i.e. a God-eyes-view. Prior could not become reconciled to the idea that divine foreknowledge could be seen as founded on what humans would choose to do in the future. Consequently, there is no truth now about an agent's future decision until the decision is actually made (Øhrstrøm 2009, p. 23, [22]). To Prior, it seemed reasonable that we should consider some possible futures, and only those propositions, which hold in all possible futures, i.e., are necessarily true in the future, can be said to be true now. In Prior's view, indeterminism is reflected in the asymmetry between the past as necessary, which means that $P(x)p \vee P(x) \sim p$ ('in the past, it has been the case at time x that either p or not p was the case') is a theorem, as precisely one of the propositions will be true of the past. On the other side, the Ockhamist theorem $F(x)p \vee F(x) \sim p$ ('in the future it is going to be the case at time x that either p or not p') should not be ascribed a truth-value now (Jakobsen, Øhrstrøm and Hasle 2017, [7]). Consequently, in the Peircean system, the future equals the necessary future. Obviously, this position runs counter to our intuition about time. As such, the Ockhamist system seems a good reflection of our common sense intuition about time, as we, in the present, act, plan and reason about the future envisioning that a certain courses of events is going to play out. By the same token, the philosophical notion of an indeterministic worldview can be seen as closely related to the idea of future contingents, represented by future statements, which are neither unavoidable nor impossible. In our context, the notion of a branching time model may serve as a highly useful framework for illustrating ways in which moral agents respond with ethical excellence by reacting at the right time, in the right way, and to the right extent, i.e., by "hitting the kairos" (Eikeland 2008, p. 154, [6]).

3 The Kaioros of Phronetic Enactment

The Aristotelian concept of phronesis reflects the kind of practical wisdom or situated knowledge which is needed for our choice and performance of morally good actions cases. Hence, in the *Nicomachean Ethics*, Aristotle mentions two kinds of virtues, viz. *intellectual* virtues (phronesis, techne, nous, episteme, and sophia) and *ethical* virtues, virtues of character. *Intellectual* virtues are further specified by distinguishing between those virtues which deal with practical thinking, versus those which theoretical reasoning (Aristotle 1934: NE, Book II., [2]).

That is, *ethical* virtues reflect standards of good practice informed by rational reasoning. As MacIntyre remarks, "According to Aristotle then excellence of character and intelligence cannot be separated" (MacIntyre 2000, p. 154, [9]). This stands in contrast to the modern worldview, according to which can be both good and stupid, as when Kant holds that possesion of the good will "[...] alone is both necessary and sufficient for moral worth" (MacIntyre 2000, p. 154, [9]). Moreowever, Aristotle holds that a flourishing life is sustained by activities, which are enacted by reason and are in alignment with virtue or excellence (Dunne 1993, p. 275, [5]). In that sense, Aristotle consolidates the superiority of the excercise of theoretical reasoning and contemplation. At first glance, this might seem to contradict his overall practice-based position and his observation that there cannot be a theoretical catalogue of rules for ethically wise decision making, since we are dealing with particular cases, which calls for excercise of practical wisdom and ethical excellence. In a similar veine, Aristotle emphasizes that young people may be good at mathematics and geometry, but that lack of practical knowledge makes will tend to make poor when it comes to moral choices:

> [...] although the young may be experts in geometry and mathematics and similar branches of knowledge, we do not consider that a young man can have Prudence. The reason is that Prudence includes a knowledge of particular facts, and this is derived from experience, which a young man does not possess; for experience is the fruit of years.
>
> (Aristotle 1934: NE, Book VI., viii. 3-5, p. 349-51, [2])

Nontheless, Aristotle distinguishes between *intellectual* virtues, obtained via teaching, and virtues of *character* and *ethics*, which are developed

via habitual exercise (MacIntyre 2000, p. 154, [9]). Here, phronesis provides the phronimos with the ability to reason well about particular cases and to see what is needed by relying on particular as well as general knowledge (nous). As such, phronesis is necessary and sufficient for being a virtuous person, who knows how to live well (eu zén). Accordingly, the phronimos is already virtous, and phronesis may inform his deliberations. Consequently, phronesis stands out from other intellectual virtues by being interweawed with the *ethical* virtues; indeed, in some interpretations of Aristotle, phronesis is considered an intellectual virtue of a special kind because it is, at the same time, an ethical virtue (Eikeland 2008, p. 53, [6]). However, as shall be discussed below, although life-experience is necessary for virtue and phronesis, it is not sufficient, and it is not the case that experience automatically fosters phronesis (Eikeland 2008, p. 154, [6]).

In continuation of this overall outline of phronesis, the role of time, viz., the kairos of phronetic activities may be elaborated on by turning to an analysis of the act of promising. Obviously, the phronimos is capable of "hitting the kairos, or the "extremely correct" target of "ethical excellence in the situation" (Eikeland 2008, p. 154, [6]). But not everyone acts with ethical excellence, when making promises. Sometimes stakes are too high, as in the well-known case of Jephthah's dilemma:

> Then the spirit of the Lord came upon Jephthah, and he passed through Gilead and Manasseh and passed on to Mizpah of Gilead, and from Mizpah of Gilead he passed on to the Ammonites. And Jephthah made a vow to the Lord, and said "if thou wilt give the Ammonites into my hand, then whoever comes forth from the doors of my house to meet me, when I return victorious from the Ammonites, shall be the Lord's, and I will offer him up for a burnt offering.
>
> (Judges 11: 29-40, NRSV, BibleGateway, [3])

Here, upon his return, Jephthah finds himself in a dilemma as he is met by his daughter, who he then has a plight to offer in sacrifice. Consequently, Jephthah can be said to have promised to do the forbidden, as argued by von Wright: "It may be shown that if the act of an agent gives rise to conflicting duties, then this act is itself something from which the agent has a duty to abstain" (Wright 1981, p. 119, [20]). Or, as formulated in his "Jephthah's Theorem": (where O stands for Ought, and "$\sim$"

is negation): $O(\sim t|p) \Rightarrow O\sim$p; i.e., if, given that p, every t is forbidden, then p itself is forbidden (Wright 1968, p. 79, [21]).

However, following van Eck (1982, [18]), it does not seem plausible to evaluate the status of a deontic expression on behalf of later unforeseen circumstances – "if Jephthah had been met by a goat, then he would not have promised the forbidden" (van Eck 1982, p. 260, [18]). Moreover, when we make a promise, we are usually aware that we risk breaking it in case later conflicting duties arise. As such, we calculate with a *ceteris paribus* clause, and we are ready to accept a certain amount of risk. Nevertheless, we feel that Jephthah's disposition is wrong; not because he has promised to do the forbidden, instead, by stating *this* specific promise at *that* given time, he has run an unusually great risk. Hence, according to van Eck, it makes sense to judge the moral status of a promise relatively to the situation in which one finds oneself, while promising to do something (van Eck 1982, p. 261, [18]). As mentioned earlier, the notion of kairos reflects the opportune moment of ethical action. This rich and complex concept has been described by Smith by pointing out three perspectives on kairos:

> Thus kairos means the time when something should happen or be done, the 'right' or 'best' time; it means the time when a constellation of events presents a crisis to which a response must be made; it means the time when an opportunity is given for creative action or for achieving some special result that is possible only at 'this' time. Implicit in all three senses of kairos is the concept of an individual time having a critical, ordinal position. Such a time is one marked off from its predecessors and successors by the fact that it is a critical or 'turning' point in a historical process.
>
> (Smith 1969, p. 6, [16])

This sense of the right time, or the favorable moment, to react to a certain situation can be reflected in a branching time model, which allows for an asymmetric account of time. Hence, Thomason (1981, [17]) presents a tempo modal model supported by two deontic principles for evaluation, depending on whether we find ourselves in a *deliberative* or a *judgmental* context. The first-mentioned refers to moral choice situations in which it is assumed that we have morally valid options, viz. alternative possible futures can be realized, and we have the opportunity to

make morally good choices. For instance, the following expression is true at nine o'clock: 'I have to go shopping before eleven o'clock'. Here, 'ought' implies 'can', but after eleven o'clock, the expression is false as it represents an option which cannot be realized any longer. Here, we may introduce a distinction between primary versus secondary, and actual duties. After having failed to do the shopping, I now have to choose among a new set of future alternatives (secondary duties). Accordingly, possible future scenarios, which are morally acceptable, could include, e.g., apologizing (a repairing obligation). However, the expression may still be regarded as valid after eleven o'clock, namely according to a *judgmental* context in which we take departure in scenarios as "might-have-beens" (Thomason 1981, p. 174, [17]), i.e., the fact that at a certain past point in time, before eleven o'clock, there existed possible alternative scenarios in which I might have acted as I ought to do.

Thomason's notion of branching time is useful to illustrate how awareness of ways in which time may both widen and constrain our freedom of choice, is significant to phronetic activities. Similarly, when we make a promise, we are aware that this involves a certain amount of risk-taking on our side since the act of promising limits our space of liberty of action in the future, and, furthermore, we may fail to succeed to do that which we ought to do. Also, when we evaluate moral actions, those of our own or others, we pass judgment by applying counterfactual reasoning, moving back in time to a point in which options were still open to us. From this vantage point, we may judge a given course of events, and in doing so, we often gain insights on contemporary events by reflecting on what we could not have known back then, either because we lacked knowledge, or because we could not, in the midst of events, know how things were to develop. Accordingly, through life-experience, we learn to master kairos, i.e., to moderate and balance moral deliberation about particular cases on the backdrop of our common-sense notion of time. As mentioned above, ethical excellence and virtue are cultivated via life experience and habituation - "they establish themselves as "skill-patterns" emerging from actions repeated many time" (Eikeland 2008, p. 182, [6]). Obviously, moral self-cultivation does not take place in isolation; rather, interactions with the surroundings, family, schools, and communities of practices scaffold our moral development. Here, Macintyre emphasizes that "the making and sustaining of those relationships is inseparable from the development of those dispositions and activi-

ties through which each is directed towards becoming an independent practical reasoner" (Macintyre 2000, p. 107, [9]).

> It is by taking part in transactions with our fellow-men that some of us become just and others unjust; by acting in dangerous situations and forming a habit of fear or of confidence we become courageous or cowardly. [...] In a word, our moral dispositions are formed as a result of the corresponding activities.
>
> (Aristotle 1934: NE, Book II. i. 7-ii. 3, p. 73-75, [2])

4 Machine ethics based on deep learning, and why we should worry

When we deliberate about how what we choose to do now may either narrow or widen our future possibilities, we are aware that there is something at stake. This kind of awareness fuels the engagement needed in our striven-towards moral self-cultivation and towards becoming phronimos. We acquire knowledge of particulars from participating in practice, dealing with cases when we, or others, have undertaken obligations, made commitments, or deliberated about matters. In such situations, more often than not, ethical insights are painfully acquired. Gradually, through experiences, we learn (often the hard way) to master the kairos of phronetic activities. In that sense, an artificial ethical agent shall presumably never be phronimos, as it would carry out ethical reasoning without the ability to gain a proper understanding of the role of time in moral life. Hence, an artificial moral agent would likely give a treatment of time by weighting the probability of different possible future outcomes on behalf of present patterns in data, patterns, about which it does not, or cannot, care. Consequently, it would lack the ability to experience the relation between time and risk, and the role this relation plays in moral life. This does not imply that machine ethics might never be realised. Presumably, a system might be capable of ethical decision making by reasoning about data sets representing both particulars and general knowledge. However, although computational ethical decision models might be feasible, they would lack an essential component of human moral agency, viz. experience based knowledge of what it means to *be* in an ethical situation in which something is at

stake. First and foremost, as pointed out by McDermott, an ethical agent is always faced with a conflict between self-interest and what ethics prescribe us to do. The fact that I may deliberately choose to act ethically wrong is constitutive of being an ethical agent (McDermott 2011, p. 95, [10]). Moreover, the fact that our learning process is driven by there being something at stake is crucial to fostering the kind of engagement needed to cultivate us to become phronimos.

But, why should we be concerned about lack of inner states in a system, which might be able to produce ethically correct judgments? It seems to be of little relevance, especially if an artificial ethical agent eventually becomes better than us in exercising ethical excellence. However, this aspect is highly relevant as machine-learning technology increasingly powers our world. Here, computational models based on deep learning methods, including supervised or unsupervised learning, have succesfully made progress in a variety of fields, such as natural language processing, topic classification, question answering, image recognition, and transcription of speech to text. Originally, conventional machine-learning techniques required engineering by hand in the shape of:

> [...] careful engineering and considerable domain expertise to design the feature extractor that transformed the raw data [...] into a suitable internal representation [...] from which the learning subsystem [...] could detect or classify patterns in the input. (LeCun et al. 2015, p. 436, [8])

Yet, recent progress in machine learning holds promise that systems will increasingly be better at learning by doing without human oversight. Hence, Lecun et al. point to "representation learning" as a set of methods, which makes it possible for a machine to automatically discover the representations needed for detection or classification (LeCun et al., 2015, p. 436, [8]). Accordingly, nowadays, humans are increasingly moving out of the loop, and this is welcomed by the authors:

> We think that deep learning will have many more successes in the future because it requires very little engineering by hand, so it can easily take advantage of increases in the amount of available computation and data.
> (LeCun et al. 2015, p. 436, [8])

Of course, the performance of such systems is measured to test "its ability to produce sensible answers on new inputs that it has never seen during training" (LeCun et al. 2015, p. 437, [8]). In that sense, it is possible to scaffold the systems responsiveness to its surroundings to avoid cases like Microsoft's fail with the real time Tay Twitter Bot, who overnight became abusive and racist as people coaxed it into saying horrible things. Nevertheless, such inductively based test methods may not guarantee an expected, and valid, outcome for all future cases (Bringsjord et al. 2006, [4]).

Moreover, with humans moving out of the loop, we do not gain an insight into how the artificial moral agent made the decision. This concern is addressed in the European General Data protection legislation (GDPR).[2] Here, transparency is one of the cornerstones in the GDPR, which takes black-box decisions into considerations (e.g., article 4, 13, 14, 22):

> The data subject shall have the right to obtain from the controller confirmation as to whether or not personal data concerning him or her are being processed, and, where that is the case, access to the personal data and the following information: [...] (h) the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.
>
> (GPDR, Article 15)

However, despite the good intentions behind the law, one may fear that it might be a difficult or even impossible task to explain, in ordinary language, what goes on due to the complexity of autonomous decision making. It is highly questionably that a deep learning driven artificial ethical agent will be able to produce an understandable output, which explains its decision making procedure in complex cases with moral ambiguity.

[2]GDPR:https://eur-lex.europa.eu/eli/reg/2016/679/oj. Accessed March 13 2019.

5 Concluding Remarks

In this paper, I have argued that discussions about whether phronesis is computationally tractable or not may benefit from including an analysis of the relation between time and morality. Hence, I have illustrated how Prior's philosophy of time and the notion of branching time may serve as a useful framework for sorting out the relation between time, free choice, and ethics. Moreover, I have shown that when we learn to master the kairos of phronetic enactment this learning process is driven by our awareness that something is at stake. Consequently, an artificial moral agent shall presumably never be able to be phronimos in the strong sense as it would have no understanding of the crucial role of time in our moral life. Hence, if we increasingly trust artificial moral machines based on complex abstractions without humans in, or on, the loop, we risk losing sight of what ethics is about in the first place. We risk switching ethics for calculation.

Bibliography

[1] Anscombe G.E.M. (1967). Aristotle And The Sea Battle. In: Moravcsik, J.M.E. (ed.) *Aristotle. Modern Studies in Philosophy.* London: Palgrave Macmillan.

[2] Aristotle, (1934). *Nicomachean Ethics*. Rackham, H. (trans.). Cambridge: MA. London: Harvard University Press, William Heinemann Ltd.

[3] BibleGateway, NSVR, Judges book. https://www.biblegateway.com/passage/?search=Judges+11%3A29-40&version=NRSV;RSV Accessed Nov 12, 2018.

[4] Bringsjord, S. Arkoudas, K. and Bello, P. (2006). Towards a General Logicist Methodology for Engineering Ethically Correct Robots. *IEEE Intelligent Systems* **21**(4): 38–44.

[5] Dunne, J. (1993). *Back to the Rough Ground – Practical Judgment and the Lure of Technique*. Notre Dame, Indiana: University of Notre Dame Press.

[6] Eikeland, O. (2008).*The Ways of Aristotle – Aristotelian Phrónêsis, aristotelian philosophy of dialogue, and action research*. Bern. Switzerland: Peter Lang AG, International Academic Publishers.

[7] Jakobsen, D., Øhrstrøm, P. and Hasle, P. (2017). In Celebration of Past, Present and Future. In: Hasle, P. Blackburn, P. and Øhrstrøm, P. (eds.): *Logic and Philosophy of Time – Themes from Prior*, Vol. 1, Aalborg University Press. Logic and Philosophy of Time Volume 1: pp. 9–27.

[8] LeCun, Y., Bengio, Y. and Hinton, G. (2015). Deep Learning. *Nature* **521**(28): 436–444.

[9] MacIntyre, A. (2000). *After Virtue – a study in moral theory*. Duckworth, London: Oxford University Press.

[10] McDermott, D. (2011). What Matters to a Machine?. In: Anderson, M. and Anderson, S. (eds.): *Machine Ethics*. New York: Cambridge University Press, New York: pp. 88–114.

[11] Prior, A.N. (1966). Postulate-sets for Tense-logic. In: Malpas, A. and Øhrstrøm, P. (eds): *The Nachlass of A.N. Prior* http://research.prior.aau.dk/cms/uploads/pdf/proofread/20111031214519.pdf. Accessed Nov 13, 2018.

[12] Prior, A.N. (1954). The Paradoxes of Derived Obligation, *Mind* **63**(249): 64–65.

[13] Prior, A.N. (1957). *Time and Modality*. Oxford University Press.

[14] Prior, A.N. (1958). Some Free Thinking about Time. In: Øhrstrøm, P. (ed.): *The Nacklass of A.N. Prior*. http://research.prior.aau.dk/nachlass/item.php?id=20120113201359. Accessed Nov 13, 2018.

[15] Prior, A.N. (1967). *Past, Present and Future*. Oxford: Clarendon Press.

[16] Smith, J.E. (1969). TIME, TIMES AND THE 'RIGHT TIME'; "CHRONOS" AND "KAIROS". *The Monist* **53**(1): 1–13.

[17] Thomason, R. (1981). Deontic Logic as Founded on Tense Logic. In: R. Hilpinen (ed.) *New Studies in Deontic Logic*. Reidel Publishing Company, Dordrecht, Holland: pp. 165–176.

[18] van Eck, J. A. (1982). A System of Temporally Relative Modal and Deontic Predicate Logic and its Philosophical Applications. *Logique et Analysis* **25**(99): 249–290.

[19] Wright, G.H. von, (1956). A note on deontic logic and derived obligation, *Mind* **65**(260): 507–509.

[20] Wright, G. H. von, (1981 reprint). A new System of Deontic Logic. In: Hilpinen, R. (ed.): *Deontic Logic: Introductory and Systematic readings*. Dordrecht: Reidel: pp. 105–121.

[21] Wright, G. H. von (1968). *An Essay in Deontic Logic and the General Theory of Action*. Amsterdam: North-Holland.
[22] Øhrstrøm, P. (2009). In Defence of the Thin Red Line: A Case for Ockhamism. *HUMANA.MENTE - Journal of Philosophical Studies* **8**: 17–32.
[23] Øhrstrøm, P., Schärfe, H. and Ploug, T. (2010). Branching Time as a Conceptual Structure. *Conceptual Structures: From Information to Intelligence*. 18th International Conference on Conceptual Structures, ICCS 2010, Kuching, Sarawak, Malaysia, July 26–30, 2010. Proceedings. LNAI, volume 6208
[24] Øhrstrøm, P. and Hasle, P. (2011). Future Contingents. In: Zalta, E.N. (ed.): *The Standford Encyclopedia of Philosophy.* https://plato.stanford.edu/entries/future-contingents/. Accessed Nov 13, 2018.

The Prior eArchive as Virtual Research Environment: towards Serendipity and Explorability

Fatima Sabir
University of Copenhagen, Denmark
fatima-sabir@hotmail.com

Volkmar Paul Engerer
University of Copenhagen, Denmark
volkmar.engerer@hum.ku.dk

Abstract

The Prior Virtual Lab is an online research collaboration which aims to transcribe and make public manuscripts, notes and letters by (or sent to) the logician and philosopher Arthur N. Prior (1914-1969). This article explores and discusses the Danish Prior Internet Resources (PIR), which the Prior Virtual Lab is part of, and argues that it constitutes a Virtual Research Environment (VRE). The article also argues that a significant challenge for future work with the PIR is applying the knowledge of the iHumanist. In particular, we shall argue that part of what is required when moving into the digital humanities is a shift from a needy-user paradigm (emphasizing search for missing information) to an explorability paradigm (which emphasizes serendipity).

Keywords: Virtual Research Environments, Serendipity, Explorability, Needy-user paradigm, eArchive, iHumanist, eResearch, Interdisciplinary Research Collaboration, Research Infrastructure, Arthur Prior, Digital Humanities.

1 Introduction

In 2011, Norbert Loussau predicted that "virtual research environments will establish themselves as the norm in the coming five to ten years, and become as entrenched as the use of email and internet in the everyday life of a researcher." (Lossau 2011, p. 156, [31]) While the prediction has three more years to be fulfilled, virtual research environments (henceforth VREs) are a long a way from becoming the norm. However, interest in VREs has not diminished, and a significant amount of research today takes place as eResearch (Jeffery and Wusteman 2012, [26]), where researchers explore and make use of modern technologies to facilitate research, research communication and research collaboration. Loussau encourages librarians to play a role in the development of eResearch by arguing:

> Transcending the borders between disciplines, librarians will be able to apply themselves to their genuine tasks of cataloguing, administering and safekeeping the accessibility of knowledge in the age of internet and at the same time take on an essential role in the research process. »Librarians Go Research!« could become the catchy motto for libraries in regard to virtual research environments and should not be frightening but, rather, an incentive for bringing our basic competencies and experience to the new digital world of collaborative research, which is new to many researchers and librarians alike. (Lossau 2011, p. 156, [31])

As information professionals (the term was explored in Engerer and Sabir, 2018, [17]), we strive to play such a role in the Danish-based eResearch project *The Primacy of Tense: A.N. Prior Now and Then* (often shortened to the *Prior Project*). Here we support an interdisciplinary group of domain researchers[1] in their use of eResearch tools and their endeavour to transform the physical Nachlass of Prior into an enhanced *virtual* Nachlass, thus moving from a traditional archive to an eArchive.

[1] "Domain researcher" is a term from information science that refers to colleagues from other disciplines [24]. In the Prior Project case this means the logicians, philosophers and historians, that the information professionals collaborate with and support with general expertise in digital information systems.

This article discusses the ongoing work of developing the domain group's main tool for eResearch: the Prior Internet Resources (PIR). Based on a discussion of how to define VREs, we argue that VREs are characterized by being online systems that support and facilitate research communication and research collaboration, data management and analysis, project management, and that functions via a participation architecture. We will argue that the PIR has many of the elements that characterize VREs, but not all of them, and that it has them to different degrees. However, we argue that this does not pose a problem for the characterization of the PIR as a VRE, as the list should not be evaluated as a collection of necessary characteristics, but merely as characteristics a VRE might have. Crucially, only the characteristics that are relevant for the particular domain group that the VRE is supposed to serve should be included. Moreover, we shall argue that the typical characteristics of VREs that the PIR lacks should be evaluated as potential developmental issues, not as disqualifications.

Due to our experience of keeping the PIR updated, we have come to believe that discussions of VREs and the PIR would benefit from an iHumanist-approach, where the discussions and developmental work are enhanced by the concepts and ideas of the digital humanities. Here we argue that VREs and other online systems for supporting and facilitating eResearch should undergo a shift from what we shall call a *needy-user paradigm*, where information professionals support the domain groups in their *search* for information, to an *explorability paradigm*, where information professionals support the domain group in *exploring* information. Here we argue that *serendipity* is a key concept in developing modern eAchieves: adequate online systems for eResearch like VREs should not just offer search based information, but should facilitate serendipitous research.

We proceed as follows. Section 2 presents the Prior Project. Here we argue that the project has enabled the Nachlass material, left behind by the New Zealand philosopher and logician, Arthur N. Prior, to move from a physical archive into an eArchive. In Section 3, we discuss the concept of VREs and present a list of characteristics a VRE typically has. Against this background, we argue that the PIR constitutes a VRE. We also argue that by making using of the knowledge of information professionals with humanistic backgrounds, *iHumanists*, the PIR has the potential to be enhanced and expanded. In Sections 4 and 5, we develop

these ideas further. In Section 4, we argue for a shift from the needy-user paradigm to an explorability paradigm. In Section 5, we argue that VREs such as the PIR could be significantly enhanced by emphasising an explorability paradigm with a focus on serendipity. Based on readings of selected literature, we present a tentative list of design features for explorable systems that foster serendipitous discoveries. Section 6 concludes.

2 From archive to eArchive

The Primacy of Tense: A.N. Prior Now and Then was launched in the autumn of 2016 as an interdisciplinary research collaboration with two different, but connected, goals. One was to explore the link between the ideas of the New Zealand born logician and philosopher Arthur N. Prior (1914-1969) and modern temporal, modal, and hybrid logic. Another was to study and expand *The Nachlass of A.N. Prior*, thereby making more of Prior's work accessible not only for researchers, but for anyone interested in Prior's logic and philosophy (see the project description at https://www.prior.aau.dk/). A significant background resource supporting these endeavours is the Danish Prior Internet Resources (PIR) consisting of the following sites:

Name	**URL**
Prior Studies	https://priorstudies.org/
The Virtual Lab for Prior Studies	https://research.prior.aau.dk/login_user.php
The Nachlass of A.N. Prior	https://research.prior.aau.dk/nachlass/

The first site, *Prior Studies*, offers traditional resources such as bibliographical lists and information about the Prior archives at Aalborg University and Oxford. The site also links to the YouTube channel *PriorStudies*, where talks on Prior's logic and philosophy can be found.

Since the spring of 2018, the PIR has been undergoing reconstruction: the design has been updated, making the site more user-friendly, and new features of have been added (the site was originally launched in the nineties, and several features needed updating).

One expansion of the research resources was the addition of a list of web resources users might find relevant for their studies. A more

notable addition to the site was the development of a research community, where researchers are invited to join a Prior-community and receive news relevant for their study of Prior.

In addition, *Prior Studies* functions as a gateway for exploring the archives at Aalborg University, the Bodleian Library, and Oxford University. One element of the gateway is information about the content of the material contained at these physical Prior archives. Another element of the gateway is the link to *The Nachlass of A.N. Prior*. Here researchers may find transcriptions of physical content from the Bodleian Library and from Aalborg and Oxford Universities.[2] At the time of writing, *The Nachlass of A.N. Prior* contained transcriptions of 68 previously unpublished letters and papers from the Prior archives. Furthermore, *Prior Studies* links to *The Virtual Lab for Prior Studies*, where photographs of the material in Oxford (more than 6000 individual photographs) may be accessed. Presently, only a limited amount of the material at Aalborg University is to be found in the Virtual Lab, but more is being added. The establishment of *The Virtual Lab for Prior Studies* was a significant contribution to research on Prior, and it was officially inaugurated at the Arthur Prior Centenary Conference held in 2014 in Oxford. This timely event was the culmination of several earlier events aimed at preserving Prior's legacy.

In 2007, Prior's widow, Mary Prior (1922-2011) granted Peter Øhrstrøm (Aalborg University) and Per Hasle (then at Aalborg University, now University of Copenhagen) permission to make use of and publish archival material from Prior located in Oxford.

The archival material was received shortly after the death of Prior in 1969, after Mary Prior and the philosopher Peter Geach (a friend of Prior) had sorted and organized notes, scientific material and letters written or received by Prior into twenty-nine boxes. These archival boxes were handed over to the Bodleian Library, where the material underwent a second ordering at the hands of two Oxford librarians, David and Steffi Lewis. From then on, researchers with written permission from the Prior family have been able to access and study the material of twenty-two boxes in the Bodleian's Special Collections Reading Room,

[2]We will describe the transcription process shortly. Here we will simply remark that the result of the transcriptions process (the files which are accessible in The Nachlass of A.N. Prior) are searchable PDF files. For a discussion of what they call the *transcription loop* see [16].

while seven boxes can be studied at the Philosophy Department Library, 10 Merton Street, Oxford.

Then, in 2010, the Bodleian Library allowed the Danish research group for Prior Studies to store digital photographs of the boxes' contents on the condition that the library retained the rights to the photographs and that access to them was to be suitably restricted. This second condition was meet by the creation of the online research platform *The Virtual Lab for Prior Studies*. Via this portal, the photographs of Prior's archival material could be accessed from anywhere in the world, but only by users who had been assigned a user profile and login. Once registered, users could download the photograph material, but were not permitted to distribute it.

Further development of *The Virtual Lab for Prior Studies* took place in 2017, when the son of Arthur and Mary Prior, Martin Prior, donated a collection of letters, notes, scrapbook material and some of Arthur Prior's personal books to Aalborg University. More recently, in late 2018 and early 2019, Martin Prior donated further material (all of these additional documents are currently in the process of being scanned and added to the lab).

However, the main purpose of *The Virtual Lab for Prior Studies* is to make the archival material accessible for examination and transcription. Registered users are free merely to browse and download the material, but the hope is that they will become active transcribers, converting this material, into searchable digital material (PDF files). It is worth noting that the transcription process makes *two* distinct changes to the accessibility to the Prior material. First, it converts it from an (unsearchable) image format to the standard format for digital documents, namely PDF files. Second, transcription transforms the ownership. The Bodleian Library retains the rights to the images, but the rights to the transcribed material reside with the Danish Prior Group.[3] All in all, transcription plays a pivotal role in the PIR, and we need to discuss it further.

When a researcher has undertaken and completed the work of transcribing a document in the Virtual Lab, another researcher (with expertise, on the one hand, in Prior's logic or his philosophy or the historical development of his work, and, on the other hand, significant experience with reading Prior's handwriting) proofreads the transcription.[4] Once

[3]The copyright holders are Per Hasle, Peter Øhrstrøm and David Jakobsen.

[4]Mary and Arthur Prior's son, Martin Prior has over the years played a significant

approved,[5] the document is uploaded to its final destination: *The Nachlass of A.N. Prior*. The path of the Prior material can be illustrated in the following way:

Figure 1: From Archive to Nachlass

Another way to think about the transformation of material is described by Engerer and Sabir:

> Here is another way to think about it. The contents of the boxes in the Bodleian Library might be thought of as Prior's 'physical Nachlass'. The photographs of this raw material in PVL [The Prior Virtual Lab[6]] might be thought of as Prior's 'virtual Nachlass'. And the transcribed and commented material available on the internet site, https://nachlass.prior.aau.dk/ (the Nachlass of A.N. Prior) might be thought of as Prior's 'digitally enriched Nachlass'.
>
> (Engerer and Sabir 2018, p. 16, [17])

Figure 2 illustrates this description of the relationship:

Figure 2: Different types of archives

In short: *The Nachlass of A.N. Prior* offers an eArchive of transcribed documents from *The Virtual Lab for Prior Studies*, which contains digital

consultative role in the transcription process, resolving numerous issues ranging from idiosyncrasies of handwriting to providing background information.

[5] Approval has to be granted by either Per Hasle, Peter Øhrstrøm or David Jakobsen.

[6] In other articles, PVL abbreviates *The Prior Virtual Lab*. However, *The Virtual Lab for Prior Studies* is the correct name.

photographs of the physical material held for the most part at Oxford. Broadly speaking the goal of the PIR is to increase accessibility to the works of Prior. Viewed from this perspective, how does it fare?

While access to both the physical Nachlass at Oxford and The Virtual Nachlass in *The Virtual Lab for Prior Studies* are restricted (the documents at the Bodleian Library can only be accessed after traveling to Oxford and obtaining a written permission from the Prior family to study the material, while the Virtual Lab requires a user profile and login) the Virtual Nachlass clearly offers more convenient access conditions. The first and the most obvious advantage is that the lab is accessible from anywhere in the world. A second difference is that the assignment of a user profile and login to the lab functions through a peer-system; probably a more transparent and open process. Third, the site makes available certain meta-information about the material stored at Oxford. These are contained in the so-called 'box-descriptions' made by Per Hasle on priorstudies.org; these offer guidelines to the content of *The Virtual Lab for Prior Studies*. These guidelines also apply to the Physical Nachlass at Oxford, and indeed they are called box-descriptions because they describe the content of the 29 physical boxes in Oxford.

The Digitally Enriched Nachlass of A.N. Prior is even more assessable, as the material in *The Nachlass of A.N. Prior* can be accessed and searched by anyone from anywhere in the world. Furthermore, since the documents in *The Nachlass of A.N. Prior* have been enriched with comments from both the transcriber and the expert who did the proofreading, the documents offer more guidance to the study of Prior's logic and philosophy. Thus, the PIR (constituted by *Prior Studies*, *The Virtual Lab for Prior Studies*, and *The Nachlass of A.N. Prior*) does fulfil its basic functions in enhancing the opportunities to study Prior.

But can its potential be enhanced and expanded? In the next section, we discuss the PIR as a Virtual Research Environment, and argue that useful development paths may be partially identified by viewing it through the lens of VRE theory.

3 On Being a Virtual Research Environment

The arrival of the personal computer, followed by their interlinking, and then, a decade later, the widespread use of the internet, transformed research, research collaboration and research communication. Before

this, research interaction tended to be limited to local environments, and was only rapid in limited contexts, such as face-to-face discussion at conferences and workshops.[7] Today, research has the potential to be both fast and global. With tools like email, video conferencing, online writing platforms, and discussion forums, research communication and research collaboration have substantially changed, so much so that it is tempting and plausible to rethink research as eResearch (Jeffery and Wusteman 2012, [26]).

However, the internet, the personal computer, and the tools they provide not only alter research communication and research collaboration, they alter the nature of research itself. Under these new conditions for research, Ankeney and Leonelli have argued that the Kuhnian idea that scientific change takes place as scientific problems accumulate (Kuhn 1962, [28]) does not offer an adequate picture of how contemporary research functions. Scientific changes, indeed scientific revolutions, are not merely the result of researchers and their scientific findings, but are also (and perhaps more importantly) the result of changed conditions for conducting research. These changed conditions influence and alter the scientific questions and problems that researchers ask and try to solve (Ankeny and Leonelli 2016, [1]).

Labels for the new transforming conditions, all designed to embrace the nature of eResearch, include 'e-research communities', 'collaborative virtual environments', 'gateways', 'science gateways', 'portals', 'virtual organizations', 'virtual research communities', 'cyber-environments' and 'virtual research environments'. Although these labels differ in meaning, definition, and the particular technologies they inhabit, they all seem to point towards the same phenomenon: digitalization of research (Bracken et al. 2014, [7]; Carusi and Reimer 2010, [9]). However, as evolving concepts, and with many different technological manifestations, the definitions of technological solutions which enable and enhance eResearch, are subject to considerable debate (Bracken et al., 2014, [7]; Carusi and Reimer, 2010, [9]; Jeffery and Wusteman 2012, [26]).

[7] The history of Prior's own research demonstrates this. Living in New Zealand, attending academic conferences was not easy, and many of Prior's research discussions with other scholars took place by letter. For example, the Smart-correspondences shows how research topics were discussed, and in one of his letters to Prior J.J.C. Smart writes, "You've missed a few things out, surprising as it may seem!", and then goes on to list events in academia, that might have interested Prior [36].

In 2010 the UK Joint Information Systems Committee (JISC) suggested VREs as significant for eResearch.[8] In their report, JISC argued that a VRE is a technological, internet-based support system, which enables eResearch to take place while potentially crossing boundaries of disciplines and locations. In their search for a definition, JISC offered the following:

> The term VRE is now best thought of as shorthand for the tools and technologies needed by researchers to do their research, interact with other researchers (who may come from different disciplines, institutions or even countries) and to make use of resources and technical infrastructures available both locally and nationally. The term VRE also incorporates the context in which those tools and technologies are used. The detailed design of a VRE will depend on many factors including discipline, context, and security requirements.
>
> (JISC in: Carusi and Reimer 2010, p. 13, [9])

That is: the key element of VREs is technological support of research in a way that has the potential to cross the boundaries of discipline and geography. Given the wide scope of this definition, determining what is and what is not a VRE can be challenging. Here, Jeffery and Wusteman have argued that the concept of VRE risks being diluted by too wide an application due to a tendency to call any online research facilitation a VRE:

> [...] as the concept of the VRE becomes more widely recognised, the tendency to describe every portal, gateway and

[8] From the many different labels for technologies that support eResearch we have chosen to discuss the PIR solely as a virtual research environment, even though other labels might have been relevant. However, as the JISC report and Jeffery and Wusteman argue, the different labels overlap significantly, and VREs stand out as a broad and flexible concept (they have many different features, not all of which need be incorporated in any particular VRE). Moreover, eResearch is frequently discussed in the context of VREs. The emphasis on multiple many different features and the non-excluding nature of the VRE-concept makes it useful for discussing the PIR – we need a flexible concept. We also remark that much discussion of eResearch takes place in a natural or social science setting, and not all features of eResearch here are relevant for eResearch in the humanities. A flexible concept helps us to transfer and open the discussion of eResearch in a humanistic setting.

> digital library as a VRE needs to be guarded against. Whilst these applications may be central components of a VRE, the latter is more than a digital library, or even a portal or gateway to a range of digital libraries. A VRE should describe an environment in which research collaboration is facilitated, not just a resource to be used in research.
>
> (Jeffery and Wusteman 2012, p. 135, [26])

In the light of this, Jeffery and Wusteman argue that the system they explore (OJAX++) is "demonstrating the next generation of Virtual Research Environments." (Jeffery and Wusteman 2012, part of the title of their article, [26]). They argue that OJAX++ facilitates research in "sharing of data; support for communication within a team; provision of access to tools; services or an infrastructure; support for project management; collaborative annotation of data; and analysis and processing of data." (Jeffery and Wusteman 2012, pp. 135–36, [26]). Jeffery and Wusteman further add that OJAX++ as a VRE offers "data and services that facilitate "remixing" with other data and services, [and functions] via an "architecture of participation"" (Jeffery and Wusteman 2012, p. 135, [26]).

Examining Jeffery and Wusteman (henceforth J&W) and the definition of VRE by the Joint Information Systems Committee (JISC), we find the following formulations, which indicate characteristic elements of VREs:

1. "Tools and technologies needed by researchers to do their research" (JISC)

2. "interact with other researchers (who may come from different disciplines, institutions or even countries)" (JISC)

3. "make use of resources and technical infrastructures available both locally and nationally" (JISC)

4. "incorporates the context in which these tools and technologies are used" (JISC)

5. "more than a digital library, or even a portal or gateway to a range of digital libraries" (J&W)

6. “describe an environment in which research collaboration is facilitated, not just a resource to be used in research” (J&W)
7. “sharing of data” (J&W)
8. “support for communication within a team” (J&W)
9. “provision of access to tools” (J&W)
10. “support project management” (J&W)
11. “collaborative annotation of data” (J&W)
12. “analysis and processing of data” (J&W)
13. “data and services that facilitate “remixing” with other data and services” (J&W)
14. “architecture of participation” (J&W)

This rather long “double” list gives grounds for characterizing VREs as online systems containing many different features relevant for the conduct of modern research. However, in some cases merely keeping to the list above, or more precisely, keeping to the exact wording, would give an oversimplified picture when we wish to discuss an online system. For example, the first item on the list gives too broad a definition of VRE: “Tools and technologies needed by researchers to do their research” could refer to just about anything from blackboard and chalk to computers and spaceships. Likewise, item 2, which requires that the system enables researchers to “interact with other researchers (who may come from different disciplines, institutions or even countries)”, could be fulfilled by email. Conversely, simply sending an email clearly does not count as making use of a virtual research environment; a key element of a VRE is indeed research communication (item 8), but also research collaboration, which is why email alone is not sufficient.

When discussing a system as a VRE, or arguing why a system perhaps does not constitute a VRE, one should keep in mind that merely holding one of the features of the list is not sufficient to qualify as a VRE (which is why simply offering a digital library, or guiding researchers on to other websites, does not qualify as a VRE (item 5)). Only when

a system incorporates a selection of the items, does it qualify as a virtual environment for eResearch. The key is that the different elements *together* facilitate research (item 6), and thus the system must offer ways of managing and analysing data (item 7, 9, 11, 12 and 13). Furthermore, the system should support project management (item 10).

Note that the third item on the list suggest that a VRE should allow research to take place 'both locally and nationally'. Here we remark that changing this to locally, nationally *and globally* seems better, as VREs are online systems that can be used anywhere.

The fifth point (VREs being more than just digital libraries or gateways) means that if we were merely looking at priorstudies.org and its function as a gateway to *The Virtual Lab for Prior Studies* and *The Nachlass of A.N. Prior*, we would *not* be dealing with a VRE. However, we shall argue below that, as a whole, the PIR does much more than simply provide these points of entry.

Being global and transnational is all very well and good, but the fourth item from the list demands that VREs also be sensitive to the context of the research. Locality is also crucial. We interpret this as meaning that VRE must be sensitive to the particular needs of the domain researchers using the VRE.

Finally, VREs operate with an "architecture of participation" (item 14). For the PIR, the drive to enter the lab is to study novel Prior material, and the motivation to undertake the work of transcription seems to be stimulated by being credited as a transcriber when, after peer review, the document goes from The Virtual Lab for Prior Studies into The Nachlass of A.N. Prior (we will say more about this later in the article).

Thus, in the interest of simplicity and compactness, we propose to condense the previous "double" list down to the following items. A VRE is an:

1. ***Online System***, which facilitates
2. ***Research communication and research collaboration***,
3. ***Data management and analysis***, and
4. ***Project management***, that works through a
5. ***Participation architecture***.

That is, we argue that a VRE should facilitate and integrate tools for research communication and collaboration and in doing so it should offer technologies for data management and analysis, be online accessible and have a participation structure. This rather broad definition leaves sufficient room different VREs to display different features depending on the particular domain of eResearch.

Taken individually, the three different domains of the PIR do not qualify as VREs. However, taken together the PIR does exhibit many of the features that characterizing VREs. We will argue that the fact that not all features are included, and some are only built-in to a limited extent does not disqualify the PIR as a VRE. The "level of fulfilment" should not be seen as binary yes or no, but be evaluated as a scale. Furthermore, in the article "Information Professionals meet Arthur Prior" (2018, [17]) it was argued that it is important for information professionals supporting domain groups to take an iHumanist approach, to explore the particular needs of the domain group, and to develop the online support system to reflect these needs. Bearing this in mind, let us now consider in more detail why the PIR should be considered a VRE.

Let us start with item (a) from our five item summary list. Obviously, the PIR is an online system and thus has the first characteristic of a VRE: the PIR can be accessed globally. That said, it is worth noticing that the different domains of the PIR vary in openness (as discussed in the previous section). While both *Prior Studies* and *Nachlass* are publicly accessible, *The Virtual Lab for Prior Studies* is restricted to assigned users. This somewhat mutes the goal of enhancing access to material on Prior's logic and philosophy.

Now for item (b). Research communication and research collaboration are an important part of *The Virtual Lab for Prior Studies*, where researchers ask for input, and comment on and discuss transcriptions. An example of this is the interaction shown in Table 1 between two researchers, where a user of *The Virtual Lab for Prior Studies* points out a minor (but significant) mistake to the transcriber.

The example shows how researchers use the lab to facilitate discussion and thereby collaborate on gaining knowledge from the archival material. However, there are some obvious limitations: neither participants nor administrators are notified when a question is posed, nor when it is answered. Which is also demonstrated by the example by the a big time lag between the initial comment and the response.

Table 1: Correspondence from The Virtual Lab for Prior Studies.

Researcher A:
– latest edit of comment: 09-01-2015 13:28:35 (GMT+1)

This letter is Sobociński's explanation of two Leśniewski's systems - Ontology and Protothetic. The unusual notation which is used here is explained in the footnotes 7 and 8. If you find it not sufficient or if you find some mistake, please inform me. I will correct it.

In this letter Sobociński used merely Polish notation.

Researcher B:
– latest edit of comment: 10-06-2018 19:19:20 (GMT+1)

The formula in the second paragraph, CfpCfNpfp, has been transcribed incorrectly. It should be CfpCfNpfq. The last symbol in the formula is q , not p.

Record no. 1157
Title: Sobocinski to Prior 16.09.1953
Fetched by user Fatima Sabir

Recently, however, *Prior Studies* has started to expand its support for a research community. Researchers and scholars are now invited to join the network by creating a profile, which provides information about their current work place and links to personal website. Such features help to identify relevant researchers and potential collaborators for studies of Prior's work. Thus, the PIR does facilitate communication and collaboration. Nonetheless, the facilities are still rather basic; for example there is no chat forum. More remains to be done here.

Item (c) from our summary list says that a VRE should offer tools for data management and analysis. Indeed, Jeffery and Wusteman argued that it should also facilitate "remixing" with different tools (Jeffery and Wusteman, 2012, p. 135, [26]). In our case, data is constituted by the digital photographs of the archival material; here *The Virtual Lab for Prior Studies* offers a structured way of accessing this material, and

The Nachlass of A.N. Prior offers digitally enriched documents from the Prior archive in Oxford and Aalborg. Also, before they enter *The Virtual Lab for Prior Studies*, researchers can freely consult information about the content of the Prior archive at the Bodleian Library using the list of box contents compiled by Per Hasle. On *Prior Studies* one finds the categorizations "Box 1-11", "Box 12-21", "Box 22", and "Box 23-29". Here researchers may find information like the following description of "Box 22":

> This box contains the full manuscript for Prior's unpublished book 'The Craft of Formal Logic' (Finished in december 1951 and submitted to Clarendon Press 6/12/1951). A full overview of its contents is given in Craft-of-logic/table of contents, and a description of its place within Prior's authorship is given at the related manuscript Items from a dictionary of formal logic/box 6.
>
> A part of the manuscript has been published as The Doctrine of Propositions and Terms, ed. by P.T. Geach and A.J.P. Kenny. University of Massachusett [sic] Press, Amherst, 1976 (1976a).
>
> Prior's work on 'The Craft of Formal Logic' is described in Kenny 1970 as well as in the introduction to 'The Doctrine of Propositions and Terms' (the latter overlaps somewhat with Kenny 1970, but naturally also adds some details). From 'The Craft of Formal Logic' the book brings Part I, chapter 1 and 2, and Part IV, chapter 1–3; these five chapters together "make up a self-contained account of the traditional doctrine of propositions and terms" (Geach and Kenny's Introduction, p. 9).
>
> The manuscript found in box 22 is a gift from J. L. Mackie (no full copy was found among Prior's own papers after his death in 1969).
>
> P. 549 and p. 600 are missing.
>
> — Per Hasle
>
> (www.priorstudies.org)

Information of this sort, describing the content of the box, provides researchers with a starting point for determining which documents in the

Virtual Lab that might be relevant to her/his research interests. In some cases, however, the information about the boxes are rather limited, like the one for "Box 20":

> 1) Scott, Dana: Advice on Modal Logic. 1968. MS 71 p. Later published in volume ed. by K. Lambert.
>
> 2) Scott, Dana: Formalizing Intensional Notions. 1968. MS 71 p.
>
> 3) Scott, Dana: The Logic of Tenses. Dec. 1965. MS 8 p.
>
> 4) Scott, Dana: A Proof of The Independence of the Continuum Hypothesis. Stanford University. 1966.
>
> 5) Segerberg, Krister, Modal Logic Based on a Three-valued Logic. MS 12 p. Published in Theoria 1967.
>
> 6) Sellars, W.,The Identity Approach to the Mind Body Problem. MS 21 p. 1963.
>
> 7) Sellars, W.,The Intentional Realism of Everett Hall. MS 22 p.
>
> 8) Sellars, W.,Theoretical Explanation. MS 17 p.
>
> 9) Sengupta, S.S., Schotch, P., Czarny,P., Relevance: A Logic and a Calculus, University of Waterloo
>
> 10) Shwayder, David, Appendix On Time and Tense, p.200-267. MS. [...]
>
> (https://www.priorstudies.org)

Thus, in the case of "Box 20", researchers merely have a list of titles to help them whether or not the material is relevant for study and transcription. As with the tools for communication and collaboration, however, the tools offered for data management and analysis are rather simple, for example it would be useful if registered users could add further comments on the items listed or perhaps remix by drawing attention to other items. But such options are not yet available.

Item (d) on our list is that VREs are tools for facilitating project management. In *The Virtual Lab for Prior Studies*; data is organized using a color system. Thus, documents that are in the process of being transcribed are marked with red; the document is locked by the transcriber (though the administrators can still access it). Data that has

been transcribed, but is still awaiting proofreading is marked yellow. Finally, documents that are both transcribed, proofread and published in *The Nachlass of A.N. Prior* are marked green in *The Virtual Lab for Prior Studies*. This very simple marking system offers gives a basic indication of progress on individual documents, but to claim that it qualify as "facilitation of project management" seems rather generous. The list of box content on *Prior Studies* also offers some overview of available data, but as has already mentioned, the level of detail and usability could be significantly improved. Data management is also part of *The Nachlass of A.N. Prior*, but only in the sense that researchers can explore transcribed documents through keyword search.

Finally, we come to item (e): that a VRE should work through a participation architecture, meaning that the system should facilitate user contribution. In the case of the PIR, the architecture of participation is built into *The Virtual Lab for Prior Studies*, and one of the incentives it offers is both powerful and interesting.

While one incentive to work in *The Virtual Lab for Prior Studies* and transcribe material might be the altruistic goal of making the philosophy and logic of Prior more accessible, the often time consuming work of reading and transcribing is demanding. The accrediting of researchers for this labor seem to be a key building block in the architecture of participation; so to speak, the altruistic goal is nudged by the reward of recognition.

While it is possible to merely study the material in *The Virtual Lab for Prior Studies*, the incentive to go further and undertake the work of transcribing the material is fostered by researchers being credited with the effort. An example from Nachlass is the following, where David Jakobsen and Martin Prior are credited for the transcription of "Letter from J.J.C. Smart to A.N. Prior, October 9, 1951" [36] with the help from Martin Prior in the following manner:

> Editor's note: The letter is in the Prior archive box 3 at the Bodleian Library in Oxford and has been transcribed and commented by David Jakobsen and Martin Prior.

However, another more scholarly incentive to participate (and perhaps the most powerful one) is the opportunity to discover interesting, perhaps even completely new, material on Prior, or to discover novel links between the data. That is, in our view, a key incentive to participate

in the VRE is the opportunity to explore data and make *serendipitous* discoveries. This is a point we will pursue in the following sections.

Our discussion of the PIR considered as a candidate VRE is summarized in Table 2:

Table 2: Elements of VREs in the PIR.

Elements of VREs	**PS**	**VL**	**NL**
Online System	✓	✓	✓
Research communication and research collaboration	(✓)	✓	×
Data management and analysis	(✓)	✓	(✓)
Participation architecture	×	✓	×
Project Management	×	×	×

Legend: PS = Prior Studies, VL = Virtual Lab, NL = Nachlass.
✓ = Fully implemented. (✓) = Partially implemented.
× = Not implemented.

4 Digital humanities: from searching to exploring

Now that we have isolated five key elements for VREs (they are online systems, facilitating research communication and collaboration, enabling data and project management, and fostering participation), and discussed their relevance and degree of realization in the current PIR, we turn to consider the PIR from the perspective of the digital humanities. This discussion takes as its point of departure an ongoing paradigm shift in information science: the move from *searchability* to *explorability*. The former has its roots in the what we have called the needy-user paradigm, which still largely characterizes *the pre-digital humanities* approach to VREs; the latter, *explorability*, acknowledges a more open, even experimental research landscape in digital humanities which is no longer driven by information needs and knowledge gaps alone. By shifting the focus from the searchability of digital resources to the explorability of digital objects in eResearch, tangible consequences for the

developmental goals of VREs in general and the PIR in particular are emerging.

There have been numerous attempts to define digital humanities (see, for example, the collection of definitions in Terras, Nyhan, and Vanhoutte 2013b, [38]) and much work in the area focuses on enumerating concrete digital projects and digitization initiatives, which in some way are supposed to define the area (Deegan and McCarty 2012, [14]; Flanders 2012, [18]; Hockey 2012, [25]; Warwick, Terras, and Nyhan 2012, [40]). However, this extensional strategy (defining by enumerating exemplars of a class) is not a shortcut to determining the essential property the exemplars have in common (the intensional approach to definition). Knowledge of such a property (or a set of properties) might prompt richer understandings of the digital humanities.

An early and frequently cited paper by Unsworth can be seen as a first step towards such an intensional definition (Unsworth 2013, [39]). Drawing on still earlier work by Davies Shrobe, and Szolovits (1993, [13]), Unsworth describes 'doing digital humanities' (or humanities computing as the current term was at that time) as a special practice of knowledge representation. He proposes viewing 'digital humanities' as a kind of attribute, a property all digital resources have, more or less, to a certain degree. His exemplification implicitly suggests a scaling of what it means for a digital system to be digital humanities-like. In full, his charlatanism-argument goes like this:

> [D]egree matters, and one way in which that degree can be measured is by the interactivity offered to users who wish to frame their own research questions. If there is none offered, and no interactivity, then the project is probably pure charlatanism. If it offers some (say, keyword searching), then it can be taken a bit more seriously. If it offers structured searching, a bit more so. If it offers combinatorial queries, more so. If it allows you to change parameters and values in order to produce new models, it starts to look very much like something that must be built on a thoroughgoing representation. If it lets you introduce new algorithms for calculating the outcomes of changed parameters and values, then it is extremely well designed indeed.
>
> (Unsworth 2013, p. 37, [39])

We can extract three main stages from this citation, two of which are rather straightforward (stages 1 and 2), one rather less so (stage 3):

1. Digital humanities and interactivity go together: the more interactive, the more digital humanities-like the system is.

2. Digital humanities means, among other things, searching: the more sophisticated the search options are, the more digital humanities-like the system is.

3. Digital humanities enables system manipulation (we propose another term below): the larger the researcher's possibilities are to modify the system (for example, to change parameters and introduce new algorithms), the more digital humanities-like the system is.

Interactivity (1) is a popular and comprehensive concept in information science (Borlund 2013, [6]; Ruthven and Kelly 2011, [32]), human-computer interaction (Lazar, Feng, and Hochheiser 2010, [29]), computer-mediated communication (Herring, Stein, and Virtanen 2013, [22]) and established disciplines such as psychology and sociology (Kiousis 2002, [27]). Searchability (2), however, seems to be the exclusive domain of information science (Case and Given 2016, [11]). However, compared with stages 1 and 2, stage 3 seems far more nebulous. Nonetheless, we find a further hint of what Unsworth means by it in a remark a few sentences later:

> But you see the principle implied by this scale – the more room a resource offers for the exercise of independent imagination and curiosity, the more substantially well thought-out, well designed, and well produced a resource it must be.
>
> (Unsworth 2013, p. 38, [39])

Here Unsworth makes the move from digital humanities-like systems to the digital humanist her/himself. Digital humanities is something practiced by researchers with "independent imagination and curiosity". This shift in perspective comes with a shift in focus: needs-triggered search behavior moves to the background, while in the foreground we instead have researchers engaging with a digital system, driven by exploration and curiosity to discover something unexpected.

Information science has traditionally been allied with stage 2 rather than with stage 3. The challenge for a search-centred information science (and the agenda of information professionals guided by this theory) is its reliance on the scenario of an 'information need', perpetuating the picture of the 'needy' individual who is urged to apply information seeking strategies (or seek professional assistance) to meet these needs (Batley 2005, [2]; Beghtol 1986, [4]; Borlund 2013, [6]; Case 2012, [10]; Cooper 1971, [12]; Derr 1983, [15]; Limberg, Sundin, and Talja 2012, [30]; Wilson 1981, [41]).[9] No doubt, the needy-user concept has its legitimacy and merits in many areas of information scientific investigation and practical work. It is by no means an irrelevant or inappropriate scenario: historically, the need for information and the actions we take to obtain information is a basic human condition (Sandstrom 1994, 1999, [33-34]). However, in a digital humanities context, the needy-user paradigm seems somewhat conservative: it evokes a set of traditional professional values drawn from information science. These values do not seem to adequately reflect the conditions found in eResearch.

The traditional idea of information as *information-about* (Beghtol 1986, [4]), which is linked to the concept of information systems as secondary supportive resources that are encoded in searchable knowledge organizing systems, has given way in the digital humanities to the idea of information as a digital research object in its own right. Accordingly, information systems such as research databases (Hider 2012, [23]) are no longer regarded as shortcuts to information somewhere external to the system or 'signposts' to knowledge about the outer world; they are to a higher degree viewed as explorative structures which lend themselves directly to research-related investigation.

Turning to the case of the Nachlass and Prior's digitized papers, by exploring whole networks of digital records of his handwritten works (letters, draft papers and so on) research questions can be inspired by information that exists exclusively in digital form. There may be nodes connecting several drafts of a manuscript with recurring concepts de-

[9]The focus on the needy-user and the overemphasis of searchability in system design furthers the tendency to interpret researchers' actual digital explorative activities from stage 3 as stage 2 search activities. Resulting misfits can then lead to information systems that still are primarily designed for information search, although researchers approach them practically without pronounced information needs. Therefore, we make the point in this paper that we need systems more designed for explorability, and less for searchability, in the digital humanities.

veloped in series of letters, or nodes indicating changes in Prior's philosophical concept of determinism in the light of critical feedback he received. Notions such as 'access to the letters' or 'information on this or that logical concept' are then not the only ways to address a Prior scholar's research interests in a digital context. Consequently, Prior researchers' motivations for using the PIR do not arise solely, or even primarily, from the information gaps they might have experienced. Far from being needy, they are curious, experimental, and fond of discovery. This perspective offers an alternative account of Prior researchers' motivation to becoming involved in the lengthy and difficult transcription process.

5 Explorability and serendipity

Serendipity occurs when a researcher discovers something interesting by chance; the discovery was not planned or specifically sought out (Björneborn 2017, [5]; Foster and Ellis 2014, [19]). At first glance it may seem that there is a tension between serendipity and the more systematic, anticipatory and expectant attitude of the researcher. Moreover, it may seem that the very idea of serendipity (which emphasizes the idea of pure chance) is further attenuated when purposeful research activities in digital environments (of the sort conducted in the PIR) are discussed. But these tensions seem resolvable. Indeed, the tension between purposeful explorations versus serendipitous findings resonates well with a well established characterization of serendipity as a kind of research strategy opposed to pure change discovery. This connection had already been noted in Cannon's book *The Way of an Investigator* in 1945 (Cannon 1945, [8]), as Foster and Ellis note (Foster and Ellis 2014, [19]). In the book, Cannon (referring to the well-known dictum due to Pasteur: *chance favors only the prepared mind*) identifies researchers' prepared minds as the feature which distinguishes chance discovery from serendipitous findings (Foster and Ellis 2014, [19]). In more concrete terms: the researcher observes a surprising datum, evaluates its significance, and uses it eventually in theoretical reflection and new hypotheses (Foster and Ellis 2014, [19]).

We might say that both serendipity and explorability cover central phenomena in the digital humanities, but emphasize different aspects of eResearchers: the notion of *serendipity* puts focus on the more emotional,

psychological attributes of digital humanists involved such as inventive, imaginative, curious, while *explorability* highlights the systematic, experimental nature in achieving serendipitous experiences. This interdependent relationship between serendipity and explorability highlights the combination of the *systematic-exploratory* attitude of researchers and the more accidental nature of the resulting event in which something unexpected is found. Exploring and having a serendipitous experience are not the same, but they are connected.

What makes exploring attractive for researchers is the possibility of serendipity; what makes systematic exploring particularly attractive for researchers is that the probability of achieving more than one serendipitous discovery is increased; and, last but not least, what makes the 'surprising', 'emotional' element in serendipity compatible with the planned, rational systematic exploration, is the condition of a 'prepared mind' in scientific serendipity (Foster and Ellis 2014, [19]). By connecting explorability to the notion of serendipity and exploiting the strong interdependencies between these two concepts, we can get a clearer picture what happens when researchers engage with a digital resource.

In digital humanities there is a close connection of exploring a digital resource and serendipitously encountering an interesting piece of information. It thus surprises that serendipity is not explicitly treated as a topic in three milestone publications on the digital humanities, namely the volume *Debates in The Digital Humanities* (Gold 2012, [21]), the partly retrospective reader *Defining Digital Humanities* (Terras, Nyhan, and Vanhoutte 2013a, [37]), and the most recent handbook *A New Companion to Digital Humanities* (Schreibman, Siemens, and Unsworth 2016, [35]). The potential of serendipity as an informational concept in the digital humanities does not seem to have been noted, yet alone fully explored. As far as we are aware both the link between explorability and serendipity on the one hand, and between serendipity/explorability and the digital humanities on the other (as targeted at by Unsworth) have not been previously noted.

How to build explorability features into eResearch systems in a way that encourages serendipitous discoveries is a difficult topic and one we cannot discuss in detail here. However, based on a selective literature review of influential studies of serendipity and exploration phenomena (Bawden 1986, [3]; Björneborn 2017, [5]; Foster and Ford 2003, [20]; Unsworth 2013, [39]) we find that serendipity and explorability features

are a matter of how the digital researcher interacts with a system (1, Interface & search features), how she/he moves around in it (2, Navigation features), which content she/he is confronted with at several stages of her/his interaction (3, Content presentation), what the underlying representational characteristics of information items are like (4, Metadata structure), how prepared and open the systems is for interaction with researchers with different disciplinary backgrounds (5, Interdisciplinary features), and the degree to which a researcher can manipulate system features her/himself (6, User engagement). Thus, based on this review, we tentatively point to six design categories, that may guide us towards building VREs in a way that reflects a move from a searchability-paradigm to an explorability-paradigm and enhances the chance of serendipitous discoveries:

1. ***Interface & search features*** (Bawden; Foster and Ford; Unsworth)
2. ***Navigation features*** (Björneborn)
3. ***Content presentation*** (Bawden; Björneborn)
4. ***Metadata structure*** (Bawden; Foster and Ford)
5. ***Interdisciplinary features*** (Bawden; Foster and Ford)
6. ***User engagement*** (Bawden)

This list reflects six areas that have been identified by research as being relevant for systems' explorability and hence to enhance the chances of serendipitous discoveries. These six categories are not clear-cut, nor are they exhaustive or mutually exclusive, and more work needs to be done on examining them individually and the links between them. Thus, these six categories are, beyond their heuristic value for further discussions, also hypotheses as to which system aspects have an impact on researchers' explorative behavior.

6 Conclusion

We have argued that while significant work has been conducted to enhance the Prior Internet Resources (PIR) as a resource that will foster research and expand knowledge on Prior's logic and philosophy, further developmental issues can be identified by thinking systematically about virtual research environments (VREs) and the digital humanities. Following a discussion of VRE-theory, we argued that the PIR qualifies as a VRE, but also suggested that it could be improved by incorporating better tools for research communication and research collaboration, and adding tools for project management. Furthermore, we argued that the architecture for participation could be enhanced. Here we have suggested that a fruitful way to approach this developmental issue might be to include the notions of serendipity and explorability, which we regard as central to the digital humanities. Based on a literature review of influential selected contributions to the notions of explorability and serendipity, we suggest that further developments of the PIR should focus on six areas of system design and information architecture: 1) Interface & search features, 2) Navigation features, 3) Content presentation, 4) Metadata structure, 5) Interdisciplinary features, and 6) User engagement. This list calls for further research, but for the moment we argued that it offers a tentative program for the development of an enhanced explorability and a user-centered, serendipitous system that invites Prior researchers' open, prepared, and curious minds, thus making pioneering discoveries in the eArchives of Prior more likely.

We should admit that there is a risk here. When discussing and designing VREs, Jeffery and Wusteman warn that one should be aware of the danger of trying to put *too* many features into a VRE:

> The temptation for software developers to reinvent the wheel is often overwhelming. The wish to create a "perfect" application for a perceived need can over-ride the option of adopting a good-enough, widely-supported and simpler solution. Commonly, such developments result in either a replication of existing software with very little added value or software which exhibits some added value but supports only a subset of the functionality of existing products. Academic-related software has appeared particularly prone to this syndrome.

(Jeffery and Wusteman 2012, p. 136, [26])

The fundamental problem of such an endeavour is that one risks compromising integration in the VRE; too many functions without clear and central purposes risk hindering researchers' use of the VRE. Therefore, the VRE should focus on integrating tools researchers already use. However, such an approach might also limit the VRE significantly in a way that does not enable researchers to gain the all the benefits that the world of VREs offers (Jeffery and Wusteman 2012, p. 137, [26]).

Clearly, we want to avoid this danger in the case of the PIR. We believe that thinking about its development in terms of the higher-level concepts of serendipity and explorability will help us to do so. A significant purpose of the PIR is to help researchers explore the logic and philosophy of Prior by enabling restricted, but significantly easier access relevant material through an eAchieve. Although researchers and scholars may enter *The Virtual Lab for Prior Studies* with the altruistic goal of making Priors logic and philosophy more accessible to a wider public, a significant motivation for participating in *The Virtual Lab for Prior Studies* is to discover material relevant for their own research. Like all other researchers, they enter the archive to discover new questions, ideas, answers and inspirations. In such an endeavour, serendipity seems vital and might be nurtured by an architecture that encourages explorability.

Bibliography

[1] Ankeny, R.A. and Leonelli, S. (2016). "Repertoires : A Post-Kuhnian Perspective on Scientific Change and Collaborative Research." *Studies in History and Philosophy of Science* **60**: 18–28.

[2] Batley, S. (2005). *Classification in Theory and Practice*. Chandos Information Professional Series. Oxford: Chandos Publishing.

[3] Bawden, D. (1986). Information systems and the stimulation of creativity. *Journal of Information Science* **12**(5): 203-216.

[4] Beghtol, C. (1986). "Bibliographic Classification Theory and Text Linguistics: Aboutness Analysis, Intertextuality and the Cognitive Act of Classifying Documents." *Journal of Documentation* **42**(2): 84–113.

[5] Björneborn, L. (2017). "Three Key Affordances for Serendipity: Toward a Framework Connecting Environmental and Personal Fac-

tors in Serendipitous Encounters." *Journal of Documentation* **73**(5): 1053–81.

[6] Borlund, P. (2013). "Interactive Information Retrieval: An Introduction." *Journal of Information Science Theory and Practice* **1**(3): 12–32.

[7] Bracken. F., Earls, D., Madders, C., O'Leary, F., Ronan, S., Ward, C., Tolan, P. and Wusteman, J. (2014). "The potential use of online tools for scientific collaboration by biology researchers". *Aslib Journal of Information Management* **66**(1): 13-37.

[8] Cannon, W.B. (1945). *The Way of an Investigator, a Scientist's Experiences in Medical Research*. New York.

[9] Carusi, A., and Reimer, T. (2010). *Virtual Research Environment Collaborative Landscape Study.* Joint Informations Systems Committe (JISC).

[10] Case, D.O. (2012). *Looking for Information: A Survey of Research on Information Seeking, Needs and Behavior*. 3rd ed. Bingley: Emerald.

[11] Case, D.O. and Given, L.M. (2016). *Looking for Information, a Survey of Research on Information Seeking, Needs, and Behavior*. 4th ed. Bingley, UK: Emerald.

[12] Cooper, W.S. (1971). "A Definition of Relevance for Information Retrieval." *Information Storage and Retrieval* **7**(1): 19–37.

[13] Davis, R., Shrobe, H. and Szolovits, P. (1993). "What Is a Knowledge Representation?" *AI Magazine* **14**(1): 17–33.

[14] Deegan, M. and McCarty, W. (2012). *Collaborative Research in the Digital Humanities*. Farnham, Surrey, England: Ashgate Pub.

[15] Derr, R.L. (1983). "A Conceptual Analysis of Information Need." *Information Processing & Management* **19**(5): 273–78.

[16] Engerer, V. and Albretsen, J. (2017). "The Prior Internet Resources 2017: Information Systems and Development Perspectives." In: Hasle, P., Blackburn, P. and Øhrstrøm, P. (eds.): *Logic and Philosophy of Time. Themes from Prior*. Aalborg, Denmark: Aalborg University Press: pp. 223–49.

[17] Engerer, V. and Sabir, F. (2018). "Information Professionals Meet Arthur Prior." *JOLIS - Journal of Librarianship and Information Science*: 1-18.

[18] Flanders, J. (2012). "Collaboration and Dissent: Challenges of Collaborative Standards for Digital Humanities." In: Deegan, M. and McCarty, W. (eds.): *Collaborative Research in the Digital Humanities*.

Edited by. Farnham. Surrey, England: Ashgate Pub: pp. 67–80.
[19] Foster, A. and Ellis, D. (2014). "Serendipity and Its Study." *Journal of Documentation* **70**(6): 1015–38.
[20] Foster, A. and Ford, N. (2003). "Serendipity and information seeking: an empircal study". *Journal of Documentation* **59**(3): 321–40.
[21] Gold, M.K. (2012). *Debates in the Digital Humanities*. Minneapolis, Minn: University of Minnesota Press.
[22] Herring, S.C., Stein, D. and Virtanen, T. (2013). *Pragmatics of Computer-Mediated Communication*. Handbook of Pragmatics, 9. Berlin: De Gruyter Mouton.
[23] Hider, P. (2012). *Information Resource Description: Creating and Managing Metadata*. London: Facet.
[24] Hjørland, B. (2002). "Domain Analysis in Information Science: Eleven Approaches – Traditional as Well as Innovative." *Journal of Documentation* **58**(4): 422–62.
[25] Hockey, S. (2012). "Digital Humanities in the Age of the Internet: Reaching Out to Other Communities." In: Deegan, M. and McCarty, W. (eds.): *Collaborative Research in the Digital Humanities*, edited by. Farnham, Surrey, England: Ashgate Pub: pp. 81–92.
[26] Jeffery, D. and Wusteman, J. (2012). "OJAX++: Demonstrating the next Generation of Virtual Research Environments." *The Electronic Library* **30**(1): 134–45.
[27] Kiousis, S. (2002). "Interactivity: A Concept Explication." *New Media & Society* **4**(3): 355–83.
[28] Kuhn, T.S. (1962). *The Structure of Scientific Revolutions*. 2nd ed. The University of Chicago Press.
[29] Lazar, J., Feng, J.H. and Hochheiser, H. (2010). *Research Methods in Human-Computer Interaction*. Hoboken, N.J.: John Wiley.
[30] Limberg, L., Sundin, O. and Talja, S. (2012). "Three Theoretical Perspectives on Information Literacy." *Human It* **11**: 93–130.
[31] Lossau, N. (2011). "Virtuelle Umgebungen." *Virtuelle Forschungsumgebungen und die Rolle von Bibliotheken* **58**(3–4): 154–63.
[32] Ruthven, I., and Kelly, D. (2011). *Interactive Information Seeking, Behaviour and Retrieval*. London: Facet.
[33] Sandstrom, P.E. (1994). "An Optimal Foraging Approach to Information Seeking and Use." *The Library Quarterly* **64**(4): 414–49.
[34] Sandstrom, P.E. (1999). "Scholars as Subsistence Foragers." *Bulletin of the American Society for Information Science and Technology*

25(3): 17–20.

[35] Schreibman, S., Siemens, R. and Unsworth, J. (eds.) (2016). *A New Companion to Digital Humanities*. Chichester, West Sussex, UK: John Wiley & Sons Inc.

[36] Smart, J. (1951) Letter from J.J.C. Smart to A.N. Prior, October 9, 1951. In: Jakobsen, D. and Prior, M. (transcribed): *The Nachlass of A.N. Prior*.https://research.prior.aau.dk/cms/uploads/pdf/proofread/20111026220810.pdf.

[37] Terras, M.M, Nyhan, J. and Vanhoutte, E. (2013a). *Defining Digital Humanities, a Reader*. Farnham, Surrey, England Burlington, VT: Ashgate Publishing Limited : Ashgate Publishing Company.

[38] Terras, M.M, Nyhan, J. and Vanhoutte, E. (2013b). "Introduction." In *Defining Digital Humanities, a Reader*, edited by Melissa M Terras, Julianne Nyhan, and Edward Vanhoutte, 1–10. Farnham, Surrey, England Burlington, VT: Ashgate Publishing Limited : Ashgate Publishing Company.

[39] Unsworth, J. (2013). "What Is Humanities Computing and What Is Not?" In *Defining Digital Humanities*, edited by Melissa M Terras, Julianne Nyhan, and Edward Vanhoutte. Farnham, Surrey: Ashgate: pp. 35–48.

[40] Warwick, C., Terras, M. and Nyhan, J. (2012). *Digital Humanities in Practice*. London: Facet publishing.

[41] Wilson, T.D. (1981). "On User Studies and Information Needs." *Journal of Documentation* **37**(1): 3–15.

AALBORG UNIVERSITY PRESS

This is a second volume of papers discussing the history of Arthur Prior's life and work and the philosophical themes he introduced or elaborated. Many of them draw inspiration from his rich and varied contributions to logic. The second volume draws on material presented at two conferences: one held in Copenhagen from 22nd-24th November 2017, and a shorter event held at Roskilde University on 2nd March 2018 as part of the *Annual Meeting of the Danish Philosophical Society*. Both meetings were organized as part of the Danish Research Council (DFF) project *The Primacy of Tense: A. N: Prior Now and Then*.